



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 96683

TO: Minh-Tam Davis
Location: CM1/8A01/8E12
Art Unit: 1642

6/18/03 2003

Case Serial Number: 991681

From: P. Sheppard
Location: CM1-1E03
Phone: (703) 308-4499

sheppard@uspto.gov

Search Notes

Priority date: 4/98 or 4/97?

request sub. for priority date 7/18/03

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

96683

From: Davis, Minh-Tam
Sent: Monday, June 16, 2003 3:28 PM
To: STIC-Biotech/ChemLib
Subject: Search request for 09/991681

Please search in commercial database, PGPUB and issued patent files:
SEQ ID NO:27-31.

Thank you.

MINH TAM DAVIS
ART UNIT 1642, ROOM 8A01, MB 8E12
305-2008

CRFE

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 6/18/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)


```

OY 1 RIRAAQOVFMIDTQCSPTKPNNDHOSCOLITELPDEKPNCHTKSVSFREIVSLL 60
DB 1290 RIRAAQOVFMIDTQCSPTKPNNDHOSCOLITELPDEKPNCHTKSVSFREIVSLL 1349
OY 61 SHOVLONLYDILLEEFVKGSPGEKTIQVPEAKLAGFLRTYSMOMLAVIFDILLDSYR 120
DB 1350 SHOVLONLYDILLEEFVKGSPGEKTIQVPEAKLAGFLRTYSMOMLAVIFDILLDSYR 1409
OY 121 TAREFDTSPGLKCLLKRVSGIGCANLYROSANSFNIFYHALVCAVLTNOETITAEQVK 180
DB 1410 TAREFDTSPGLKCLLKRVSGIGCANLYROSANSFNIFYHALVCAVLTNOETITAEQVK 1469
OY 181 VLEFEDERSTDSQCSSEDEDIFETTAQVSPRGKRRKQRRAMPPLSVQVSNADWV 240
DB 1470 VLEFEDERSTDSQCSSEDEDIFETTAQVSPRGKRRKQRRAMPPLSVQVSNADWV 1529
OY 241 LVRLHKLKLMELCNNTYIOHMLDLENCEBPPIFGDPFLLPSFOSSSTPTSGSGKE 300
DB 1530 LVRLHKLKLMELCNNTYIOHMLDLENCEBPPIFGDPFLLPSFOSSSTPTSGSGKE 1589
OY 301 TPSEDRSOSREHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKITYMAD 360
DB 1590 TPSEDRSOSREHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKITYMAD 1649
OY 361 KTISKLMTEYKKRKOONLSAPFKEVKEKGEPLGPRGODSPLLOPQHLMDQOMRS 420
DB 1650 KTISKLMTEYKKRKOONLSAPFKEVKEKGEPLGPRGODSPLLOPQHLMDQOMRS 1709
OY 421 FSAGPELLRQDKRPRSGSGSSLSVSRDAEAQIOAMTNMVLTVLNOIILPDQFTALQ 480
DB 1710 FSAGPELLRQDKRPRSGSGSSLSVSRDAEAQIOAMTNMVLTVLNOIILPDQFTALQ 1769
OY 481 PAVFPCISQLTCHVTDIRVQAVREMLGVRGVYDIIV 518
DB 1770 PAVFPCISQLTCHVTDIRVQAVREMLGVRGVYDIIV 1807

RESULT 2
OY 096CH9 PRELIMINARY: PRT: 592 AA.
AC 096CH9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical 66.4 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RM (1)
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014227; AAH14227.1;
KW Hypothetical protein.
FT NON-TER
SO SEQUENCE 592 AA: 66400 MW: 7418807B0DF4785 CRC64;

Query Match 99.28; Score 2675.5; DB 4; Length 592;
Best Local Similarity 99.68; Pred. No. 5.5e-217;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

DB 196 TAREFDTSPGLKCLLKRVSGIGCANLYROSANSFNIFYHALVCAVLTNOETITAEQVK 255
OY 181 VLEFEDERSTDSQCSSEDEDIFETTAQVSPRGKRRKQRRAMPPLSVQVSNADWV 240
DB 256 VLEFEDERSTDSQCSSEDEDIFETTAQVSPRGKRRKQRRAMPPLSVQVSNADWV 314
OY 241 LVRLHKLKLMELCNNTYIOHMLDLENCEBPPIFGDPFLLPSFOSSSTPTSGSGKE 300
DB 315 LVRLHKLKLMELCNNTYIOHMLDLENCEBPPIFGDPFLLPSFOSSSTPTSGSGKE 374
OY 301 TPSEDRSOSREHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKITYMAD 360
DB 375 TPSEDRSOSREHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKITYMAD 434
OY 361 KTISKLMTEYKKRKOONLSAPFKEVKEKGEPLGPRGODSPLLOPQHLMDQOMRS 420
DB 435 KTISKLMTEYKKRKOONLSAPFKEVKEKGEPLGPRGODSPLLOPQHLMDQOMRS 494
OY 421 FSAGPELLRQDKRPRSGSGSSLSVSRDAEAQIOAMTNMVLTVLNOIILPDQFTALQ 480
DB 495 FSAGPELLRQDKRPRSGSGSSLSVSRDAEAQIOAMTNMVLTVLNOIILPDQFTALQ 554
OY 481 PAVFPCISQLTCHVTDIRVQAVREMLGVRGVYDIIV 518
DB 555 PAVFPCISQLTCHVTDIRVQAVREMLGVRGVYDIIV 592

RESULT 3
OY 09W444 PRELIMINARY: PRT: 2045 AA.
AC 09W444:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE CG5937 protein.
GN CG5937.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
DE Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RM (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busan S.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gehlert W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlisina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington V.

```

THIS PAGE BLANK (USPTO)

GenCore version 5.2.0
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:51:11 ; Search time 9.95608 Seconds
(without alignments)
827.825 Million cell updates/sec

Title: US-09-991-681-30
Perfect score: 215
Sequence: 1 SPYKEDKPSRRKEMENAGNKITYMAADTKISKLTETK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organellar:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriophage:*
17: SP-archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	592	4	096CH9
2	215	100.0	1807	4	09ULH6
3	65	30.2	177	5	09BUT8
4	64	29.8	621	2	059301
5	63.5	29.5	604	3	074491
6	61	28.4	3306	10	09FT44
7	59	27.4	277	16	08U805
8	58	27.0	672	10	0941M7
9	57.5	26.7	921	16	092826
10	56.5	26.3	201	16	08XBF3
11	56.5	26.3	483	10	064951
12	56.5	26.3	574	10	064950
13	56	26.0	304	10	09SIS6
14	56	26.0	328	2	08S341
15	56	26.0	408	2	09L8P3
16	55.5	25.8	103	17	08TV49

17	55.5	25.8	676	5	09N4T5	09n4t5 caenorhabdi
18	55	25.6	177	10	024127	024127 nicotiana t
19	55	25.6	351	10	09SMC1	09smc1 nicotiana r
20	55	25.6	351	10	09SMB7	09smb7 nicotiana t
21	55	25.6	351	10	09MB29	09mb29 nicotiana t
22	55	25.6	654	10	09FLI2	09fli2 arabidopsis
23	54.5	25.3	555	10	09AV96	09av96 nicotiana t
24	54.5	25.3	673	10	094IM6	094im6 hordeum vul
25	54	25.1	732	2	08VP74	08vp74 pseudomonas
26	54	25.1	1048	10	004938	004938 oryza sativ
27	53.5	24.9	331	2	093EX6	093ex6 rhodococcus
28	53.5	24.9	1055	10	09LIY9	09liy9 oryza sativ
29	53	24.7	176	5	09BIF0	09bif0 plasmodium
30	53	24.7	186	4	09POX1	09pox1 homo sapien
31	53	24.7	259	5	093905	093905 caenorhabdi
32	53	24.7	373	10	08RUE5	08rue5 oryza sativ
33	53	24.7	410	10	09S2X1	09s2x1 arabidopsis
34	53	24.7	423	10	09FK39	09fk39 arabidopsis
35	53	24.7	462	16	007732	007732 mycobacteri
36	53	24.7	462	16	08VJ06	08vj06 mycobacteri
37	53	24.7	488	5	09VFE6	09vfe6 drosophila
38	53	24.7	528	4	09N003	09n003 homo sapien
39	53	24.7	729	5	021136	021136 caenorhabdi
40	53	24.7	900	10	09FK71	09fk71 arabidopsis
41	52.5	24.4	234	10	09S7C4	09s7c4 hevea brasl
42	52.5	24.4	769	3	08X1E4	08x1e4 neurospora
43	52.5	24.4	895	6	09BDV9	09bdv9 lepus capen
44	52.5	24.4	1289	16	09W287	09w287 thermotoga
45	52	24.2	196	11	09D068	09d068 mus musculu

ALIGNMENTS

RESULT 1	096CH9	PRELIMINARY:	PRT:	592 AA.
ID	096CH9	01-DEC-2001 (TREMBLrel. 19, Created)		
AC	096CH9	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hypothetical 66.4 kDa protein (Fragment).			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=COLON;			
RA	Stratberg R.;			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC014227; AAH14227.1;			
KW	Hypothetical protein.			
FT	NON TER			
SO	SEQUENCE 592 AA; 66400 MW; F4A1E807BD6F47B5 CRC64;			
Query Match	100.0%;	Score 215;	DB 4;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 6e-20;		
Matches	40;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
DB	406	SPYKEDKPSRRKEMENAGNKITYMAADTKISKLTETK 445		
QY	1	SPYKEDKPSRRKEMENAGNKITYMAADTKISKLTETK 40		
DB	406	SPYKEDKPSRRKEMENAGNKITYMAADTKISKLTETK 445		
RESULT 2	09ULH6	PRELIMINARY:	PRT:	1807 AA.
ID	09ULH6	01-MAY-2000 (TREMBLrel. 13, Created)		
AC	09ULH6	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	K1A1244 protein (BIG3) (Fragment).			

THIS PAGE BLANK (USPTO)

GN KIAA1244.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirotsawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RN for large proteins in vitro.";
RN Res. 6:337-345(1999).

SEQUENCE OF 38-1807 FROM N.A.
RA Hong W.;
RT "KIAA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7P
RT subfamily of ARP GEFs.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB033070; BAA6558.1;
DR EMBL: AF413080; AAL04174.1;
DR InterPro: IPR000904; Sec7.
DR SMART: SM02222; Sec7; 1.
FT NON_TER
SQ SEQUENCE 1807 AA; 199921 MW; B762C29916F72CB1 CRC64;

Query Match 100.0%; Score 215; DB 4; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2.1e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SPVKKKDSRRKKEMWNGNKITYTMAADKTTSLKMTK 40
DB 1621 SPVKKKDSRRKKEMWNGNKITYTMAADKTTSLKMTK 1660

RESULT 3
ID 09BJE8 PRELIMINARY; PRT; 177 AA.
NC 09BJE8;
JT 09-JUN-2001 (TREMBlrel. 17, Created)
JT 09-JUN-2001 (TREMBlrel. 17, Last sequence update)
JT 09-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=701;
RX MEDLINE=21136462; PubMed=11237850;
RA "Khatib A., Kun J., Deloron P., Kremsner P.G., Klunkert M.O.;
RT "Variants of Plasmodium falciparum Erythrocyte Membrane Protein 1
RT Expressed by Different Placental Parasites are Closely Related and
RT Adhere to Chondroitin Sulfate A";
RJ Infect. Dis. 183:1165-1169(2001).
R EMBL: AF334805; AAK28128.1;
FT NON_TER
T NON_TER
T NON_TER
SQ SEQUENCE 177 AA; 21207 MW; 318AC7FEF0102A2 CRC64;

Query Match 30.2%; Score 65; DB 5; Length 177;
Best Local Similarity 52.6%; Pred. No. 0.89;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
6 KQDSRRKKEMWNGNKITYTMAADKTTSLKMTK 24
112 KKDDTKRKEMWNGNKSHIN 130

SULT 4
9301
059301 PRELIMINARY; PRT; 621 AA.

AC 059301;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Endo-beta-1,4-xylosylase precursor (EC 3.2.1.8).
OS Cellvibrion mixtus.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Cellvibrion.
ON NCBI_TaxID=39650;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIXTUS;
RX MEDLINE=96077124; PubMed=7492333;
RA Millward-Sadler S.J., Davidson K., Hazlewood G.P., Black G.W.,
RA Gilbert H.J., Clarke J.H.;
RT "Novel cellulose-binding domains, NodB homologues and conserved
RT Pseudomonas fluorescens subsp. cellulosae and Cellvibrion mixtus.";
RL Biochem. J. 312:39-48(1995).
DR EMBL: Z48926; CAAB8762.1;
DR HSP; P14768; ICLX.
DR InterPro: IPR005088; CBM_15.
DR InterPro: IPR001000; Glyco_hydro_10.
DR Pfam: PF03426; CBM_15; 1.
DR PRINTS: PR00134; GLYCO_HYDROL_10; 1.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; UNKNOWN_1.
FT GLYCOSYLASE; Hydrolase; Signal; Xylan degradation.
FT SIGNAL
SQ SEQUENCE 621 AA; 64929 MW; 830E58959D2C3AB8 CRC64;

Query Match 29.8%; Score 64; DB 2; Length 621;
Best Local Similarity 43.2%; Pred. No. 4.9;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;
QY 2 PKVKKKDSRRKKEMWNGNKITYTMAADKTTSLKMTK 38
DB 175 PKVKKKDSRRKKEMWNGNKITYTMAADKTTSLKMTK 211

RESULT 5
ID 074491 PRELIMINARY; PRT; 604 AA.
AC 074491;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative ATP-dependent RNA helicase C285.03.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
ON NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Seeger K., Harris D., Lyne M., Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 115-604 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.
DR EMBL: AL031545; CAA20842.1;
DR EMBL: D89259; BAA13920.1;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000620; DEAD

THIS PAGE BLANK (USPTO)

(KRA7/) KRA7OCHVIL J D.

Human pectidase enco
Novel human diagno
Bovine mammary LLS
Prophion Bacterium
Human Immune/Haema
Drosophila melanog
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Aribidopsis thalia
Arabidopsis thalia
Encoded by human G
Human foetal proteo
Novel human diagno
Amino acid sequenc
Amino acid sequenc
Human protein SEQ
Human protein SEQ
Novel human diagno
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Mutated chymidylact
Human chymidylate
Peptide #1137 eno

PA (ROBE/) ROBERTS-RAPP L.
 XX Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-187683/24.
 DR
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 PS Claim 17; Page 44; 57pp; English.
 XX
 XX The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 393-441 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC AAM51651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 49 AA;
 Query Match 100.0%; Score 260; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. NO.1.3e-28;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPLGPRQDSPLLRPQHLMDQGMHRSFSGAPPELLRQDKRRPRSGTGS 49
 1 EPLGPRQDSPLLRPQHLMDQGMHRSFSGAPPELLRQDKRRPRSGTGS 49
 DB
 RESULT 2
 AAM5472
 ID AAM85472 standard; Protein: 518 AA.
 XX
 AC AAM85472;
 XX
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WC9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 XX (ABBO) ABBOTT LAB.
 XX
 XX Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
 PI Robert-Strapp L, Russell JC, Stroupe SD;

XX
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.
 XX
 PT New p118 nucleic acid and proteins - used for diagnosis and
 PT treatment of prostatic disease, especially cancer, and also for drug
 PT screening
 PS Claim 17; Page 93-94; 117pp; English.
 XX
 XX The present sequence is encoded by consensus PS118 sequence derived from
 CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
 CC were identified from a cDNA library made from prostate tumour tissue.
 CC Recombinant PS118 protein is used to detect PS118-specific antibodies,
 CC to raise antibodies for detection of PS118 antigens, and to screen for
 CC specific binding agents (potential therapeutics, and to isolate specific
 CC antibodies from serum. Detection of PS118 protein or nucleic acid, which
 CC are prostate related, and altered or elevated in prostatic disease, is
 CC used for detection, diagnosis, staging, monitoring and prognosis of
 CC prostatic disease, particularly cancer, and to identify subjects at
 CC risk.
 CC
 SQ Sequence 518 AA;
 Query Match 100.0%; Score 260; DB 19; Length 518;
 Best Local Similarity 100.0%; Pred. NO.2.1e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPLGPRQDSPLLRPQHLMDQGMHRSFSGAPPELLRQDKRRPRSGTGS 49
 393 EPLGPRQDSPLLRPQHLMDQGMHRSFSGAPPELLRQDKRRPRSGTGS 441
 DB
 RESULT 3
 AAM50809
 ID AAM50809 standard; Protein: 518 AA.
 XX
 AC AAM50809;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker partial sequence.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 XX (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 XX Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-187683/24.

DR N-PSDB: ABA91651.
XX
PT Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
PS Claim 17; Page 42-43; 57pp; English.
XX
XX The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus cDNA sequence
CC (see ABA91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing, and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.
XX
SQ Sequence 518 AA:
Query Match 100.0%; Score 260; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.1e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPLGPRGDSPLLPQPHLMQGMHRSFSAPELLRDKRPSGSTGS 49
DB 393 EPLGPRGDSPLLPQPHLMQGMHRSFSAPELLRDKRPSGSTGS 441
RESULT 4
ABG09728
ID ABG09728 standard; Protein: 1807 AA.
XX
AC ABG09728;
XX
XX 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9719.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS73915.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
PS Claim 20; SEQ ID No 40087; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1807 AA:
Query Match 100.0%; Score 260; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 9e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPLGPRGDSPLLPQPHLMQGMHRSFSAPELLRDKRPSGSTGS 49
DB 1682 EPLGPRGDSPLLPQPHLMQGMHRSFSAPELLRDKRPSGSTGS 1730
RESULT 5
ABG09731
ID ABG09731 standard; Protein: 1982 AA.
XX
AC ABG09731;
XX
XX 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9722.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS73918.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX

PS Claim 20; SEQ ID NO 40090; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG3037 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX

SQ Sequence 1982 AA;

Query Match 100.0%; Score 260; DB 22; Length 1982;

Best Local Similarity 100.0%; Pred. No. 1e-26;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPILGRGDSPLLRPQHLMDOGMRHSFSGAPPELLRQDKRPRSGSGS 49

DB 1796 EPILGRGDSPLLRPQHLMDOGMRHSFSGAPPELLRQDKRPRSGSGS 1844

RESULT 6

AAU32052

ID AAU32052 standard; Protein: 195 AA.

XX

AC AAU32052;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #2543.

XX

KM Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KM Immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US08656.

XX

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI: 2001-611725/70.

XX

PT Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy -

XX

PS Claim 20; Page 548-549; 765pp; English.

XX

CC The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising

CC the nucleic acids encoding the polypeptides and cells genetically

CC engineered to express them are also useful for producing the proteins.

CC The proteins are useful in genetic vaccination, testing and

CC therapy, and can be used as nutritional supplements. They may be used to

CC increase stem cell proliferation; to regulate haematopoiesis; and in

CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;

CC immune suppression and/or stimulation; as anti-inflammatory agents; and

CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid

CC sequences of novel human secreted proteins of the invention.

XX

SQ Sequence 195 AA;

Query Match 25.8%; Score 67; DB 22; Length 195;

Best Local Similarity 37.5%; Pred. No. 0.37;

Matches 18; Conservative 6; Mismatches 18; Indels 6; Gaps 2;

OY 4 GPRGDSPLLRPQHLMDOGMRHSFSGAPPELLRQDK--RPRSGSGS 49

DB 55 GKNQGEPAVYKCSHL-----VKHSQSRPSSWRQEKITRTKGSGPS 98

RESULT 7

ABBA0346

ID ABBA0346 standard; Peptide: 70 AA.

XX

AC ABB40346;

XX

DT 04-FEB-2002 (first entry)

XX

DE Peptide #7852 encoded by human foetal liver single exon probe.

XX

KM Human; foetal liver; gene expression; single exon nucleic acid probe.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00669.

XX

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI: 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver -

XX

PS Claim 27; SEQ ID NO 32981; 639pp + sequence listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the


```

XX  XX  MO200157272-A2.
PN  XX
XX  XX  09-AUG-2001.
PD  XX
XX  XX  30-JAN-2001; 2001WO-US000663.
PF  XX
XX  XX  04-FEB-2000; 2000US-0180312.
PR  XX  26-MAY-2000; 2000US-0207456.
PR  XX  30-JUN-2000; 2000US-0608408.
PR  XX  03-AUG-2000; 2000US-0632366.
PR  XX  21-SEP-2000; 2000US-0234687.
PR  XX  27-SEP-2000; 2000US-0236359.
PR  XX  04-OCT-2000; 2000GB-0024263.
XX  XX
XX  XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX  XX
XX  XX  Penn SG, Hanzel DK, Chen W, Rank DR;
XX  XX  WPI; 2001-48897/53.
XX  XX
XX  XX  Human genome-derived single exon nucleic acid probes useful for
XX  XX  analyzing gene expression in human placenta -
XX  XX
XX  XX  Claim 27; SEQ ID No 34316; 654pp; English.
XX  XX
XX  XX  The present invention relates to single exon nucleic acid probes (SENP:
XX  XX  see A11315-A157546). The present sequence is a peptide encoded by one
XX  XX  such probe. The probes are useful for producing a microarray for
XX  XX  predicting, measuring and displaying gene expression in samples derived
XX  XX  from human placenta. The probes are useful for antenatal diagnosis of
XX  XX  human genetic disorders.
XX  XX
XX  XX  Sequence 70 AA:
SQ
Query Match 24.2%; Score 63; DB 22; Length 70;
Best Local Similarity 34.5%; Pred. No. 0.4;
Matches 19; Conservative 7; Mismatches 13; Indels 16; Gaps 3;
QY 5 PRGDSPLRPOHLDQGMRHSFSA-----GP-----ELLRDKRPRSGTG 48
    1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1
DB 12 PPGQ-----QDPPLHHRGGRHSEAVHSHRPGVGRHLDQAVRQDEAGLGATG 61
AC  ABG43753;
XX
XX  19-AUG-2002 (first entry)
DT
XX
XX  Human peptide encoded by genome-derived single exon probe SEQ ID 33418.
DE
XX
XX  Human: single exon probe; asthma; lung cancer; COPD; ILD;
XX  XX  Chronic obstructive pulmonary disease; interstitial lung disease;
XX  XX  familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX  XX  tubercous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX  XX  Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX  XX  pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX  XX  pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX  XX  primary ciliary dyskinesia; pulmonary hypertension;
XX  XX  hyaline membrane disease.
XX  XX
XX  XX  Homo sapiens.
XX  XX
XX  XX  MO200186003-A2.
XX  XX
XX  XX  15-NOV-2001.
XX  XX
XX  XX  30-JAN-2001; 2001WO-US000665.
XX  XX
XX  XX  04-FEB-2000; 2000US-180312P.

```

```

PR  XX  26-MAY-2000; 2000US-207456P.
PR  XX  30-JUN-2000; 2000US-0608408.
PR  XX  03-AUG-2000; 2000US-0632366.
PR  XX  21-SEP-2000; 2000US-234687P.
PR  XX  27-SEP-2000; 2000US-236359P.
PR  XX  04-OCT-2000; 2000GB-0024263.
XX  XX
XX  XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX  XX
XX  XX  Penn SG, Hanzel DK, Chen W, Rank DR;
XX  XX  WPI; 2002-114183/15.
XX  XX
XX  XX  Spatially-addressable set of single exon nucleic acid probes, used to
XX  XX  measure gene expression in human lung samples -
XX  XX
XX  XX  Claim 27; SEQ ID No 33418; 634pp; English.
XX  XX
XX  XX  The invention relates to a spatially-addressable set of single exon
XX  XX  nucleic acid probes for measuring gene expression in a sample derived
XX  XX  from human lung comprising single exon nucleic acid probes having one of
XX  XX  12614 nucleic acid sequences mentioned in the specification, or their
XX  XX  complements or the 12387 open reading frames derived from the 12614
XX  XX  probes. Also included are a microarray comprising the novel set of
XX  XX  probes; the novel set of probes which hybridize at high stringency to a
XX  XX  nucleic acid expressed in the human lung; measuring gene expression in a
XX  XX  sample derived from human lung, comprising (a) contacting the array with
XX  XX  a collection of detectably labeled nucleic acids derived from human lung
XX  XX  mRNA, and (b) measuring the label detectably bound to each probe of
XX  XX  the array; identifying exons in a eukaryotic genome, comprising
XX  XX  (a) algorithmically predicting at least one exon from genomic sequences
XX  XX  of the eukaryote; and (b) detecting specific hybridisation of detectably
XX  XX  labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX  XX  having a fragment identical to the predicted exon, the probe is included
XX  XX  in the above mentioned microarray; assigning exons to a single gene,
XX  XX  comprising (a) identifying exons from genomic sequence by the method
XX  XX  above and (b) measuring the expression of each of the exons in several
XX  XX  tissues and/or cell types using hybridisation to a single exon
XX  XX  microarrays having a probe with the exon, where a common pattern of
XX  XX  expression of the exons in the tissues and/or cell types indicates that
XX  XX  the exons should be assigned to a single gene; a peptide comprising one
XX  XX  of 12011 sequences, mentioned in the specification, or encoded by the
XX  XX  probes/open reading frames (ORF). The probes are used for gene
XX  XX  expression analysis, and for identifying exons in a gene, particularly
XX  XX  using human lung derived mRNA and for the study of lung diseases
XX  XX  such as asthma, lung cancer, chronic obstructive pulmonary disease
XX  XX  (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX  XX  fibrosis, neurofibromatosis, tubercous sclerosis, Gaucher's disease,
XX  XX  Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX  XX  haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX  XX  pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX  XX  pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX  XX  and hyaline membrane disease. The present sequence is a peptide/protein
XX  XX  encoded by a single exon probe of the invention.
XX  XX  Note: The sequence data for this patent did not form part
XX  XX  of the printed specification, but was obtained in electronic
XX  XX  format directly from WIPO at
XX  XX  ftp.wipo.int/pub/published_pct_sequences.
XX  XX
SQ  Sequence 70 AA:
Query Match 24.2%; Score 63; DB 23; Length 70;
Best Local Similarity 34.5%; Pred. No. 0.4;
Matches 19; Conservative 7; Mismatches 13; Indels 16; Gaps 3;
QY 5 PRGDSPLRPOHLDQGMRHSFSA-----GP-----ELLRDKRPRSGTG 48
    1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1 11 1:1
DB 12 PPGQ-----QDPPLHHRGGRHSEAVHSHRPGVGRHLDQAVRQDEAGLGATG 61
AC  ABG22251
XX
XX  12-AUG-2002 (first entry)
DT
XX
XX  Human peptide encoded by genome-derived single exon probe SEQ ID 33418.
DE
XX
XX  Human: single exon probe; asthma; lung cancer; COPD; ILD;
XX  XX  Chronic obstructive pulmonary disease; interstitial lung disease;
XX  XX  familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX  XX  tubercous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX  XX  Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX  XX  pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX  XX  pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX  XX  primary ciliary dyskinesia; pulmonary hypertension;
XX  XX  hyaline membrane disease.
XX  XX
XX  XX  Homo sapiens.
XX  XX
XX  XX  MO200186003-A2.
XX  XX
XX  XX  15-NOV-2001.
XX  XX
XX  XX  30-JAN-2001; 2001WO-US000665.
XX  XX
XX  XX  04-FEB-2000; 2000US-180312P.

```

```

XX AC ABG22251;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #22242.
XX KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI: 2001-639362/73.
XX DR N-PSDB: AAS86438.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 52610; 103bp; English.
XX SX
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SX
SQ Sequence 171 AA:
Query Match 24.2%; Score 63; DB 22; Length 171;
Best Local Similarity 34.0%; Pred. No. 1.1;
Matches 16; Conservative 6; Mismatches 7; Indels 18; Gaps 2;
OY 2 PLGPGGDSPLQRPQHLMDGOMRHSFASAPPELLRQDKRPPRSSTG 48
DB 62 PTGPR-----TNKPKHLIKQGR-----GPPROPISGSTG 90

```

```

XX AC 15-MAY-2001 (first entry)
XX DT Bovine mammary tissue derived protein #48.
XX DE Bovine mammary tissue derived protein #48.
XX KW Bovine; mammary gland; cancer; tumour; angiogenesis.
XX KW Bos taurus.
XX OS WO200114553-A1.
XX PN 01-MAR-2001.
XX PD 23-AUG-2000; 2000WO-NZ00166.
XX PF 23-AUG-1999; 99US-0150330.
XX PR (GENE-) GENESIS RES & DEV CORP LTD.
XX PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
XX PI Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;
XX DR WPI: 2001-226619/23.
XX PT New polypeptides and polynucleotides encoding the polypeptides, which
XX PT are expressed in bovine mammary gland tissue, useful for stimulating
XX PT mammary gland growth or function, or inducing differentiation of milk
XX PT producing cells.
XX PS Claim 11; Page 80; 97bp; English.
XX SX
XX CC The present invention relates to proteins derived from bovine
XX CC mammary gland cells. The invention is useful for stimulating
XX CC bovine mammary gland cell growth and function, inhibiting the
XX CC growth of various mammary gland cancer cells, inhibiting
XX CC angiogenesis and vascularization of tumours, or modulating
XX CC the growth of blood vessels in a mammal.
XX SX
SQ Sequence 87 AA:
Query Match 23.8%; Score 62; DB 22; Length 87;
Best Local Similarity 39.0%; Pred. No. 0.71;
Matches 16; Conservative 4; Mismatches 17; Indels 4; Gaps 1;
OY 4 GPGGDSPLQRPQHLMDGOMRHSFASAPPELLRQDKRPPRS 44
DB 45 GKNGGEPTRVRCSHL-----VKHSQSRPSSMRQEKTPPS 81

```

```

RESULT 14
AAU66156
ID AAU66156 standard; Protein; 122 AA.
AC AAU66156;
DT 27-FEB-2002 (first entry)
DX
XX DE Propionibacterium acnes immunogenic protein #27052.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.

```

PR 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Sreiky YAM, Persing DH, Mitcham JJ, Mang SS, Bhatia A;
PI L' Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59716.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 27351; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 122 AA;

Query Match 22.7%; Score 59; DB 22; Length 122;
Best Local Similarity 29.9%; Pred. No. 2.8;
Matches 20; Conservative 9; Mismatches 20; Indels 18; Gaps 3;

OY 1 EPUG-----PRGQSPILQRPQHLM----DQGQ--MRHSAGPELLRDKRP 42
 ||||| | | |
DB 28 EPLGMRQKLVALCKPTARTLEVLGHPAHAVASSDGRNARPMPTDFPSAPEIPELGRS 87
 || || || ||
OY 43 RSGSTGS 49
 || || || ||
DB 88 RSMTTGS 94

RESULT 15
AAM84901
ID AAM84901 standard; Protein: 54 AA.
XX
AC AAM84901;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:12494.
XX
KW Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;
KW cytosolic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN MO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.

XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227102.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.

PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241826
PR	01-NOV-2000	2000US-0244617
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249266
PR	17-NOV-2000	2000US-0249265
PR	17-NOV-2000	2000US-0249267
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249299
PR	17-NOV-2000	2000US-0249300
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0251030
PR	05-DEC-2000	2000US-0251988
PR	05-DEC-2000	2000US-0256719
PR	06-DEC-2000	2000US-025479
PR	06-DEC-2000	2000US-0251855
PR	08-DEC-2000	2000US-0251866
PR	08-DEC-2000	2000US-0251869
PR	08-DEC-2000	2000US-0251889
PR	08-DEC-2000	2000US-0251990
PR	11-DEC-2000	2000US-0254097
PR	05-JAN-2001	2001US-0255978
PA	(HUMA-) HUMAN GENOME SCT INC.	
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI: 2001-483426/52.	
XX	N-PSDB; AAK57682.	
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides	

Pt useful for preventing, diagnosing and/or treating cancers and metastasis -
xx

ClaIm 11: SEQ ID NO 12494; 3071bp + Sequence Listing: English.

CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAAM82170 to AMM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAKR7694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAAM82169 represent sequences used in the exemplification of The present Invention.
xx

SQ **Sequence** **54 AA:**

Query Match Best Local Similarity Matches	22.5%; Score 58.5; 34.8% Pred. No. 1,2;	DB 22; Length 54; No. Mismatches 12; Indels 11; Gaps 3
--	--	---

OY 7 GDSDFLQRPOHLMQGOMRSFSAGPELLRODKRRP-----SGSTG 48
|||:::||::||::|||::|||::|||::|||:
Dd 1 CGKGVLTWKRAHPSPKDGHLP----GGPELV-EDPPRPRLPTGTGAOG 39

Search completed: June 17, 2003, 11:56:00
Job time : 15.1332 secs

THIS PAGE BLANK (JERRY)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:56:11 ; Search time 4.51977 Seconds
(without alignments)
318.982 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPGQDSPLLRPQHLM.....SAGPELLRQDKRPRSGSTGS 49Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PT05.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	49	US-09-065-383-31	Sequence 31, Appl
2	260	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	56.5	21.7	312	US-09-230-637-34	Sequence 34, Appl
4	56.5	21.7	313	US-09-347-878-30	Sequence 30, Appl
5	56.5	21.7	313	US-09-367-007C-39	Sequence 39, Appl
6	56	21.5	163	US-09-066-074-2	Sequence 2, Appl
7	56	21.5	163	US-08-555-912A-2	Sequence 2, Appl
8	56	21.5	163	US-09-208-804-4	Sequence 4, Appl
9	56	21.5	163	US-08-801-743-4	Sequence 4, Appl
10	55	21.2	1882	US-09-369-364A-13	Sequence 13, Appl
11	54.5	21.0	647	US-09-031-563-7	Sequence 7, Appl
12	54.5	21.0	647	US-09-392-277-7	Sequence 7, Appl
13	54.5	21.0	648	US-09-031-563-5	Sequence 5, Appl
14	54.5	21.0	648	US-09-392-277-5	Sequence 5, Appl
15	54.5	21.0	1055	US-09-031-563-27	Sequence 27, Appl
16	54.5	21.0	1055	US-09-392-277-27	Sequence 27, Appl
17	54.5	21.0	1315	US-09-031-563-2	Sequence 2, Appl
18	54.5	21.0	1315	US-09-031-563-25	Sequence 25, Appl
19	54.5	21.0	1315	US-09-293-505-10	Sequence 10, Appl
20	54.5	21.0	1315	US-09-392-277-2	Sequence 2, Appl
21	54.5	21.0	1315	US-09-392-277-25	Sequence 25, Appl
22	54.5	21.0	1433	US-08-365-486A-21	Sequence 21, Appl
23	54.5	21.0	1433	US-09-123-708-4	Sequence 4, Appl
24	54.5	21.0	1433	US-09-123-624-4	Sequence 4, Appl
25	54.5	21.0	1433	US-08-880-342-21	Sequence 21, Appl
26	54	20.8	260	US-08-558-135-7	Sequence 7, Appl
27	54	20.8	793	US-09-523-849-32	Sequence 32, Appl

28	53.5	20.6	284	2	US-08-751-233A-6	Sequence 6, Appl
29	53	20.4	454	1	US-08-166-316-2	Sequence 2, Appl
30	53	20.4	574	2	US-08-906-713-2	Sequence 2, Appl
31	52.5	20.2	284	2	US-08-751-233A-8	Sequence 8, Appl
32	52.5	20.2	1284	4	US-09-343-494-9	Sequence 9, Appl
33	51	19.6	265	4	US-09-134-001C-336	Sequence 336, Ap
34	51	19.6	485	2	US-08-749-391-2	Sequence 2, Appl
35	51	19.6	485	4	US-09-390-200-2	Sequence 2, Appl
36	51	19.6	510	4	US-09-522-217-89	Sequence 89, Appl
37	50.5	19.4	466	4	US-08-868-435-2	Sequence 2, Appl
38	50.5	19.4	466	4	US-08-744-231-2	Sequence 2, Appl
39	50.5	19.4	600	1	US-08-253-785-3	Sequence 3, Appl
40	50	19.2	330	3	US-08-851-843A-203	Sequence 203, App
41	50	19.2	330	4	US-08-974-549A-322	Sequence 322, App
42	50	19.2	330	4	US-08-854-050-203	Sequence 203, App
43	50	19.2	330	4	US-09-430-323-203	Sequence 203, App
44	50	19.2	570	4	US-08-969-046-4	Sequence 4, Appl
45	50	19.2	603	4	US-09-198-122-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 100.0%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 5.1e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGSGS 49
Db 1 EPLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGSGS 49

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065.383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842.385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35.441
REFERENCE/DOCKET NUMBER: 6084. US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 260; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.9e-29;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGSGS 49
Db 393 EPLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGSGS 441

RESULT 3
US-09-230-637-34
Sequence 34, Application US/09230637
Patent No. 6264958

GENERAL INFORMATION:
APPLICANT: Hayward, Gary
APPLICANT: Nicholas, John
APPLICANT: Hardwick, J. Marie
APPLICANT: Reitz, Marvin
TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
TITLE OF INVENTION: Associated Herpesvirus
FILE REFERENCE: 1107.78372
CURRENT APPLICATION NUMBER: US/09/230.637
CURRENT FILING DATE: 1999-11-23
PRIOR FILING DATE: 1996-07-25
PRIOR APPLICATION NUMBER: PCT US 97/12931
PRIOR FILING DATE: 1997-07-24
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
US-09-230-637-34

Query Match 21.7%; Score 56.5; DB 4; Length 312;
Best Local Similarity 34.0%; Pred. No. 3.3;
Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 2 PLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGS 46
Db 11 PLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGS 54

RESULT 4
US-09-347-878-30
Sequence 30, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347.878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 30
LENGTH: 313
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Human thymidylate synthase protein sequence
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: D00596/GenBank
US-09-347-878-30

Query Match 21.7%; Score 56.5; DB 4; Length 313;
Best Local Similarity 34.0%; Pred. No. 3.3;
Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 2 PLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGS 46
Db 12 PLGPRGDDSPLLORPQH-LMDQGMKHSFSGPELLRODKRPRSGS 55

RESULT 5
US-09-367-007C-39
Sequence 39, Application US/09367007C
Patent No. 6416987
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/US98/02145
PRIOR FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 39
LENGTH: 313
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: wild type human thymidylate synthase amino acid
US-09-367-007C-39
Query Match
Best Local Similarity 34.0%; Pred. No. 3.3:
Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 2;
21.7%; Score 56.5; DB 4; Length 313:
OY 2 PLPGGDSPLLRPQH-LMDGQMRHSFSGPELLRQDRPRS 46
DB 12 PLPPAQERDAEPHPHGLQYLQIQHILRCG---VKKDRFTGTGT 55
RESULT 6
US-09-066-074-2
Sequence 2, Application US/09066074
Patent No. 5952467
GENERAL INFORMATION:
APPLICANT: Hunter et al., Tony
TITLE OF INVENTION: NIMA INTERACTING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,074
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/555,912
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 163 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-066-074-2
Query Match
Best Local Similarity 21.5%; Score 56; DB 2; Length 163;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;
OY 4 GPGGDSPLLRPQH-LMDGQMRHSFSGPELLRQDRPRS 44
DB 45 GKNQGEPAVRVRCSHLL-----VKHSQSRPSSMNOEKITRT 81
RESULT 7
US-08-555-912A-2
Sequence 2, Application US/08555912A
Patent No. 5972697
GENERAL INFORMATION:
APPLICANT: Hunter et al., Tony
TITLE OF INVENTION: NIMA INTERACTING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,912A
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 163 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-555-912A-2
Query Match
Best Local Similarity 21.5%; Score 56; DB 2; Length 163;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;
OY 4 GPGGDSPLLRPQH-LMDGQMRHSFSGPELLRQDRPRS 44
DB 45 GKNQGEPAVRVRCSHLL-----VKHSQSRPSSMNOEKITRT 81
RESULT 8
US-09-208-804-4
Sequence 4, Application US/09208804
Patent No. 6030826
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
TITLE OF INVENTION: NOVEL HUMAN PARVULIN-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA

```

: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/208,804
: FILING DATE: Herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/801,743
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0217 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 163 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 1332710
: US-09-208-804-4

Query Match      21.5%; Score 56; DB 3; Length 163;
Best Local Similarity 36.6%; Pred. No. 1.8;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;

QY 4 GPRGDSPLLRPOHMDGOMRHSFSGPELLRQDKRPRS 44
Db 45 GKNGGEPARVRCSHL---VKHSQSRPRSSWNRQEKITRT 81

RESULT 9
US-08-801-743-4
: Sequence 4, Application US/08801743
: Patent No. 6037164
: GENERAL INFORMATION:
: APPLICANT: Au-Young, Janice
: TITLE OF INVENTION: NOVEL HUMAN PARVULIN-LIKE PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/801,743
: FILING DATE: Herewith
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0217 US
```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 163 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 1332710
: US-08-801-743-4

Query Match      21.5%; Score 56; DB 3; Length 163;
Best Local Similarity 36.6%; Pred. No. 1.8;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;

QY 4 GPRGDSPLLRPOHMDGOMRHSFSGPELLRQDKRPRS 44
Db 45 GKNGGEPARVRCSHL---VKHSQSRPRSSWNRQEKITRT 81

RESULT 10
US-09-369-364A-13
: Sequence 13, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hutsainen, Taina L.
: APPLICANT: Hirohata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: CURRENT FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 13
: LENGTH: 1882
: TYPE: PRT
: ORGANISM: Homo sapiens ADAMTS-9
: FEATURE:
: NAME/KEY: MOD_RES
: LOCATION: (466)
: OTHER INFORMATION: Xaa = C
: NAME/KEY: MOD_RES
: LOCATION: (521)
: OTHER INFORMATION: Xaa = Y
: US-09-369-364A-13

Query Match      21.2%; Score 55; DB 4; Length 1882;
Best Local Similarity 35.9%; Pred. No. 48;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;

QY 8 QDSPLLRPOHMDGOMRHSFSGPELLRQDKRPRS 46
Db 1235 QDCSMSPCPORTPDGSLAQHPFQ-----NEDYRPRSAS 1267

RESULT 11
US-09-031-563-7
: Sequence 7, Application US/09031563A
: Patent No. 6022708
: GENERAL INFORMATION:
: APPLICANT: Frederic de Sauvage
: APPLICANT: Arnon Roshenthal
: TITLE OF INVENTION: Fused
: FILE REFERENCE: P1272
: CURRENT APPLICATION NUMBER: US/09/031,563A
: CURRENT FILING DATE: 1998-02-26
: NUMBER OF SEQ ID NOS: 27
: SEQ ID NO 7
```

LENGTH: 647
TYPE: PRT
ORGANISM: Homo sapiens
US-09-031-563-7

Query Match 21.0%; Score 54.5; DB 3; Length 647;
Best Local Similarity 26.8%; Pred. No. 15;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

OY 3 LGPRGDSPLL-QRPQHLMDGOMRHSFSAGPELLRODKR-----PRSGST 47
DB 292 LAPKGNOSRILTOAYKRMABEAMOKKHONTGPALOEODKTSKVAAPTAPLPRIGAT 347

RESULT 12
US-09-392-277-7
Sequence 7, Application US/09392277A
Patent No. 6451977
GENERAL INFORMATION:
APPLICANT: Frederic de Sauvage
APPLICANT: Arnon Rosenthal
TITLE OF INVENTION: Fused
FILE REFERENCE: P1272R1P1
CURRENT APPLICATION NUMBER: US/09/392,277A
CURRENT FILING DATE: 1999-09-03
EARLIER APPLICATION NUMBER: US 09/258,000
EARLIER FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: US 60/076,072
EARLIER FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 28
SEQ ID NO 7
LENGTH: 647
TYPE: PRT
ORGANISM: Homo sapiens
US-09-392-277-7

Query Match 21.0%; Score 54.5; DB 4; Length 647;
Best Local Similarity 26.8%; Pred. No. 15;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

OY 3 LGPRGDSPLL-QRPQHLMDGOMRHSFSAGPELLRODKR-----PRSGST 47
DB 292 LAPKGNOSRILTOAYKRMABEAMOKKHONTGPALOEODKTSKVAAPTAPLPRIGAT 347

RESULT 13
US-09-031-563-5
Sequence 5, Application US/09031563A
Patent No. 6022708
GENERAL INFORMATION:
APPLICANT: Frederic de Sauvage
APPLICANT: Arnon Rosenthal
TITLE OF INVENTION: Fused
FILE REFERENCE: P1272
CURRENT APPLICATION NUMBER: US/09/031,563A
CURRENT FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 27
SEQ ID NO 5
LENGTH: 648
TYPE: PRT
ORGANISM: Homo sapiens
US-09-031-563-5

Query Match 21.0%; Score 54.5; DB 3; Length 648;
Best Local Similarity 26.8%; Pred. No. 15;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

OY 3 LGPRGDSPLL-QRPQHLMDGOMRHSFSAGPELLRODKR-----PRSGST 47
DB 292 LAPKGNOSRILTOAYKRMABEAMOKKHONTGPALOEODKTSKVAAPTAPLPRIGAT 347

RESULT 14

US-09-392-277-5
Sequence 5, Application US/09392277A
Patent No. 6451977
GENERAL INFORMATION:
APPLICANT: Frederic de Sauvage
APPLICANT: Arnon Rosenthal
TITLE OF INVENTION: Fused
FILE REFERENCE: P1272R1P1
CURRENT APPLICATION NUMBER: US/09/392,277A
CURRENT FILING DATE: 1999-09-03
EARLIER APPLICATION NUMBER: US 09/258,000
EARLIER FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: US 60/076,072
EARLIER FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 28
SEQ ID NO 5
LENGTH: 648
TYPE: PRT
ORGANISM: Homo sapiens
US-09-392-277-5

Query Match 21.0%; Score 54.5; DB 4; Length 648;
Best Local Similarity 26.8%; Pred. No. 15;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

OY 3 LGPRGDSPLL-QRPQHLMDGOMRHSFSAGPELLRODKR-----PRSGST 47
DB 292 LAPKGNOSRILTOAYKRMABEAMOKKHONTGPALOEODKTSKVAAPTAPLPRIGAT 347

RESULT 15
US-09-031-563-27
Sequence 27, Application US/09031563A
Patent No. 6022708
GENERAL INFORMATION:
APPLICANT: Frederic de Sauvage
APPLICANT: Arnon Rosenthal
TITLE OF INVENTION: Fused
FILE REFERENCE: P1272
CURRENT APPLICATION NUMBER: US/09/031,563A
CURRENT FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 27
SEQ ID NO 27
LENGTH: 1055
TYPE: PRT
ORGANISM: Homo sapiens
US-09-031-563-27

Query Match 21.0%; Score 54.5; DB 3; Length 1055;
Best Local Similarity 26.8%; Pred. No. 28;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

OY 3 LGPRGDSPLL-QRPQHLMDGOMRHSFSAGPELLRODKR-----PRSGST 47
DB 32 LAPKGNOSRILTOAYKRMABEAMOKKHONTGPALOEODKTSKVAAPTAPLPRIGAT 87

Search completed: June 17, 2003, 12:03:41
Job time : 5.51977 secs

THIS PAGE BLANK (use 70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:36 ; Search time 5.23719 Seconds
(without alignments)
1000.321 Million cell updates/sec

Title: US-09-991-681-31
Perfect score: 260
Sequence: 1 EPLGPRGQSPLLQRPQHLM.....SAGPELLRDKRPRSGTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	24.2	70	10 US-09-864-761-44434	Sequence 44434, A
2	56	21.5	70	10 US-09-864-761-46074	Sequence 46074, A
3	55	21.2	32	10 US-09-864-761-35805	Sequence 35805, A
4	55	21.2	1629	10 US-09-972-467-2	Sequence 2, App1
5	55	21.2	1882	10 US-09-918-171A-13	Sequence 13, App1
6	55	21.2	1907	10 US-09-938-330-25	Sequence 25, App1
7	54.5	21.0	1315	9 US-09-990-046-10	Sequence 10, App1
8	54.5	21.0	1433	9 US-10-224-249-14	Sequence 14, App1
9	54	20.8	623	9 US-10-108-605-125	Sequence 125, App1
10	54	20.8	623	9 US-10-108-605-129	Sequence 129, App1
11	54	20.8	638	9 US-10-059-585-8	Sequence 8, App1
12	54	20.8	779	1 US-08-817-832B-31	Sequence 31, App1
13	54	20.8	793	9 US-10-195-101-32	Sequence 32, App1
14	54	20.8	795	9 US-10-142-356-9	Sequence 9, App1
15	54	20.8	795	10 US-09-919-585-12	Sequence 12, App1
16	53.5	20.6	244	9 US-09-734-329-5	Sequence 5, App1
17	53.5	20.6	428	9 US-09-734-329-2	Sequence 2, App1
18	53	20.4	90	10 US-09-867-550-1392	Sequence 1392, App1
19	53	20.4	323	9 US-09-912-672A-8	Sequence 8, App1

20	53	20.4	538	9	US-09-976-740-43	Sequence 43, App1
21	53	20.4	538	12	US-10-023-529-43	Sequence 43, App1
22	53	20.4	538	12	US-10-023-523-43	Sequence 43, App1
23	53	20.4	560	9	US-09-912-672A-5	Sequence 5, App1
24	53	20.4	574	9	US-09-912-672A-2	Sequence 2, App1
25	53	20.4	574	9	US-10-063-547-164	Sequence 164, App1
26	53	20.4	574	9	US-10-063-516-164	Sequence 164, App1
27	53	20.4	574	9	US-10-063-516-164	Sequence 164, App1
28	53	20.4	574	9	US-10-063-518-164	Sequence 164, App1
29	53	20.4	574	9	US-10-063-598-164	Sequence 164, App1
30	53	20.4	574	9	US-10-227-693-164	Sequence 164, App1
31	53	20.4	574	9	US-10-063-538-164	Sequence 164, App1
32	53	20.4	574	9	US-10-063-538-164	Sequence 164, App1
33	53	20.4	574	9	US-10-063-538-164	Sequence 164, App1
34	53	20.4	574	9	US-10-063-538-164	Sequence 164, App1
35	53	20.4	574	9	US-10-223-873A-5	Sequence 25, App1
36	53	20.4	574	9	US-10-063-595-164	Sequence 164, App1
37	53	20.4	574	9	US-10-063-595-164	Sequence 164, App1
38	53	20.4	574	9	US-10-238-565-4	Sequence 4, App1
39	53	20.4	574	9	US-10-063-580-164	Sequence 164, App1
40	53	20.4	574	10	US-09-728-911-25	Sequence 25, App1
41	53	20.4	574	10	US-09-870-574-4	Sequence 4, App1
42	53	20.4	574	12	US-10-006-867-164	Sequence 164, App1
43	53	20.4	596	9	US-10-213-590-51	Sequence 51, App1
44	52.5	20.2	1284	9	US-10-160-224-9	Sequence 9, App1
45	52.5	20.2	1284	9	US-10-185-867-11	Sequence 11, App1

ALIGNMENTS

RESULT 1
US-09-864-761-44434
Sequence 44434, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OR INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemolica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44434
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004952.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.84
US-09-864-761-44434
```

```

Query Match      24.2%  Score 63; DB 10; Length 70;
Best Local Similarity 34.5%  Pred. No. 0.74; Mismatches 13; Indels 16; Gaps 3;
Matches 19; Conservative 7;
```

```

QY      5  PRGQDSPLLRQHMLDQGMHRSFSA-----GP-----ELLRDOKRRSGSTG 48
Db      12  PRCQ-----QQPHLHNRGGRHSEAVHSHRPGVRINHQAVRQDEAGLGATG 61
```

```

RESULT 2
US-09-864-761-46074
; Sequence 46074, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/632,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46074
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008752.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EST HUMAN HIT: BF314489.1, EVALUE 2.00e-28
; OTHER INFORMATION: SWISSPROT HIT: Q13526, EVALUE 2.00e-29
US-09-864-761-46074
```

```

Query Match      21.5%  Score 56; DB 10; Length 70;
Best Local Similarity 36.6%  Pred. No. 5.5; Mismatches 15; Conservative 5; Indels 17; Gaps 1;
Matches 15; Conservative 5;
```

```

QY      4  GPRGDSPLLRQHMLDQGMHRSFSAPELLLRDOKRRPS 44
Db      25  GRNGGEPARVRCSHL---VKNQSRRPSRWREKIRTT 61
```

```

RESULT 3
US-09-864-761-35805
; Sequence 35805, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
```



```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35805
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007688.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA.. SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALU0 3.90e-01
US-09-864-761-35805
```

```

Query Match          21.2%; Score 55; DB 10; Length 32;
Best Local Similarity 71.4%; Pred. No. 3;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 25 MRHSFSGAPPELLRQ 38
Db 11 LQHSFSGAPPELLQ 24
```

```

RESULT 4
US-09-972-467-2
; Sequence 2, Application US/09972467
; Patent No. US20020090373a1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; FILE REFERENCE: PC10850A
; CURRENT APPLICATION NUMBER: US/09/972,467
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1629
; TYPE: PRT
; ORGANISM: Human
US-09-972-467-2
```

```

Query Match          21.2%; Score 55; DB 10; Length 1629;
Best Local Similarity 35.9%; Pred. No. 2.7e+02;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;
```

```

QY 8 QDSPLLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGS 46
Db 1288 QDCSMSPCQRTPTDGLAQHPFQ-----NEDYRPRSAS 1320
```

```

RESULT 5
US-09-918-171A-13
```

```

; Sequence 13, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1882
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = Cys
; NAME/KEY: MOD_RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Tyr
US-09-918-171A-13
```

```

Query Match          21.2%; Score 55; DB 10; Length 1882;
Best Local Similarity 35.9%; Pred. No. 3.1e+02;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;
```

```

QY 8 QDSPLLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGS 46
Db 1235 QDCSMSPCQRTPTDGLAQHPFQ-----NEDYRPRSAS 1267
```

```

RESULT 6
US-09-938-330-25
; Sequence 25, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Halibun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: NO. US20020115838A1e1 Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-25
```

```

Query Match          21.2%; Score 55; DB 10; Length 1907;
Best Local Similarity 35.9%; Pred. No. 3.2e+02;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;
```

```

QY 8 QDSPLLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGS 46
Db 1260 QDCSMSPCQRTPTDGLAQHPFQ-----NEDYRPRSAS 1292
```

```

RESULT 7
```

```
US-09-990-046-10
; Sequence 10, Application US/09990046
; Patent No. US20020156245A1
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/990,046
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-990-046-10

Query Match      21.0%; Score 54.5; DB 9; Length 1315;
Best Local Similarity 26.8%; Pred. No. 2.4e+02;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY      3 LGPRGDSPLL-QRPHLMDQGMRSFSAGPELLRQDKR-----PRSGST 47
DB      292 LARKGQSRLTGATYRMAEAMQKKHQTGPALEQEDKTSKVAPGTAPLRILGAT 347

RESULT 8
US-10-224-249-14
; Sequence 14, Application US/10224249
; Publication No. US20030087867A1
; GENERAL INFORMATION:
; APPLICANT: Vogels, Ronald V.
; APPLICANT: Verilinder, Stefan F.F.
; TITLE OF INVENTION: Gene therapy for enhancing and/or inducing angiogenesis
; FILE REFERENCE: 2183-5233US
; CURRENT APPLICATION NUMBER: US/10/224,249
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: PCT/NL00/00482
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: EP 99202263.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/143,101
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(1433)
; OTHER INFORMATION: Human nitric oxide synthase
US-10-224-249-14

Query Match      21.0%; Score 54.5; DB 9; Length 1433;
Best Local Similarity 29.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 5; Mismatches 16; Indels 23; Gaps 3;

QY      1 EPLG-----PRGDSPLL-----QRPHLMDQGMRSFS---SAGPELLR 37
DB      124 QPLGPTKAVDLSHPKGEQPLAVDAGSGPENGQPHAYDDQEQAGSLPHANGWPGAPR 183

QY      38 QD 39
DB      184 QD 185

RESULT 9
US-10-108-605-125
; Sequence 125, Application US/10108605
```

```
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 125
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-125

Query Match      20.8%; Score 54; DB 9; Length 623;
Best Local Similarity 35.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 4; Mismatches 19; Indels 8; Gaps 2;

QY      8 QDSPLLRPOHL-----MDQGMRSFSAGPELLRQDKRPRSGSTGS 49
DB      110 QQSLLLRPOHQOQOQSHOSQOQOQHGYGSSAQL--PHNRLSGSGSTGS 155

RESULT 10
US-10-108-605-129
; Sequence 129, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-129

Query Match      20.8%; Score 54; DB 9; Length 623;
Best Local Similarity 35.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 4; Mismatches 19; Indels 8; Gaps 2;

QY      8 QDSPLLRPOHL-----MDQGMRSFSAGPELLRQDKRPRSGSTGS 49
DB      110 QQSLLLRPOHQOQOQSHOSQOQOQHGYGSSAQL--PHNRLSGSGSTGS 155

RESULT 11
US-10-059-585-8
; Sequence 8, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
```

```

APPLICANT: Isogai, Takao
APPLICANT: Mishikawa, Tetsuo
APPLICANT: Hayashi, Koji
APPLICANT: Otsuka, Kaoru
APPLICANT: Yamamoto, Jun-ichi
APPLICANT: Ishii, Shizuko
APPLICANT: Sugiyama, Tomoyasu
APPLICANT: Wakamatsu, Ai
APPLICANT: Nagai, Keiichi
APPLICANT: Otsuki, Tetsuji
APPLICANT: Funahashi, Shin-ichi
APPLICANT: Senoo, Chiaki
APPLICANT: Nezu, Jun-ichi
TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
FILE REFERENCE: 06501-098001
CURRENT APPLICATION NUMBER: US/10/059,585
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: PCT/JP00/05060
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: US 60/183,322
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: US 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 11-248036
PRIOR FILING DATE: 1999-07-29
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 638
TYPE: PRT
ORGANISM: Homo sapiens
: S-10-059-585-8

```

```

:      PRIOR APPLICATION DATA:      EP 94 11 7122.5
:      APPLICATION NUMBER:
:      FILING DATE: 28-OCT-1994
:      INFORMATION FOR SEQ ID NO: 31:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 779 amino acids
:      TYPE: amino acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: protein
:      US-08-817-832B-31
:
Query Match      20.8%; Score 54; DB 1; Length 779;
Best Local Similarity 36.2%; Pred. No. 1.5e+02;
Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2;
OY      13 LQRRQHL-----MDQGQMHRFSAGPEL---LRQDKRRPSRSGTSGS;49
db      388 LQSAHLAKVORTTISANOKRRESDHAGSITPPAVSYTKRPQANVES 434

```

```

RESULT 13
US-10-195-101-32
; Sequence 32, Application US/10195101
; Publication No. US20030087317A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, David M.
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009-1 CIP
; CURRENT APPLICATION NUMBER: US/10/195,101
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 09/523,849
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank Accession No: US20030087317A1 g2052189
; DATABASE ENTRY DATE: 25 April 1997
US-10-195-101-32

Query Match      20.8%, Score 54; DB 9; Length 793;
Best Local Similarity 36.2%; Pred. No. 1.6e+02;
Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2

Cy      13 LORPQH-----MDGOMRHSFSAGPEL---LRQDKRPPSGSGTGS 49
Db      401 LQSPHLKVORSTISNQKRRFSDHAGPSIPPAVSYTKRPQANSVES 447

RESULT 14
US-10-142-356-9
; Sequence 9, Application US/10142356
; Publication No. US20030036183A1
; GENERAL INFORMATION:
; APPLICANT: Boylan, John F.
; APPLICANT: Bowers, Alex J.
; TITLE OF INVENTION: Serine-Threonine Kinase Member, h2520-40
; FILE REFERENCE: 01017/37177A
; CURRENT APPLICATION NUMBER: US/10/142,356
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,276
; PRIOR FILING DATE: 2001-05-10

```

NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 795
TYPE: PRT
ORGANISM: Homo sapiens
US-10-142-356-9

Query Match 20.8%; Score 54; DB 9; Length 795;
Best Local Similarity 36.2%; Pred. No. 1.6e+02;
Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2;

QY 13 LQRPQHL-----MDGQMRHSFSAGPEL---LRDDKRRPSGSGTGS 49
DB 401 LQSPAHLYKVSISANQKRRFSDHAGPSIPAVSYTKRPPQANSVES 447

RESULT 15

US-09-919-585-12
Sequence 12, Application US/09919585
Patent No. US20020115167A1
GENERAL INFORMATION:
APPLICANT: Sun, Tian-Qiang
APPLICANT: Feng, Jia-Jia
APPLICANT: Reinhard, Christoph
APPLICANT: Rantl, Wendy J.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR-1
TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTILIZING
FILE REFERENCE: PP-016093.002/200130.525
CURRENT APPLICATION NUMBER: US/09/919,585
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 795
TYPE: PRT
ORGANISM: Homo sapiens
US-09-919-585-12

Query Match 20.8%; Score 54; DB 10; Length 795;
Best Local Similarity 36.2%; Pred. No. 1.6e+02;
Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2;

QY 13 LQRPQHL-----MDGQMRHSFSAGPEL---LRDDKRRPSGSGTGS 49
DB 401 LQSPAHLYKVSISANQKRRFSDHAGPSIPAVSYTKRPPQANSVES 447

Search completed: June 17, 2003, 12:02:32
Job time : 6.23719 secs

QY 5 PRGDSPLQRPHIMDOGMRSF--SAGDELLRODKRPR 43
|||:::||:::||::|||
Db 103 PDGADTILIGEDALLDDGKIRITTFVNTAGNHI-----RPR 138

Best Local Similarity 36.6%; Pred. No. 15;
Matches 15; Conservative 8; Mismatches 11; Indels 7; Gaps 2.

RESULT 12
S69336
arylsulfatase (EC 3.1.6.1) - *Pseudomonas aeruginosa*
C:Snacfas; Pseudomonas aeruginosa

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Mar-2000
C:Accession: S69336, S69359
R:Bell, S.; Kerrill, H.; James, P.; Staudenmann, W.; Cook, A.M.; Letsinger, T.; Korteaz, Eur. J. Biochem. 229, 385-394, 1995
A>Title: Purification and characterization of the arylsulfatase synthesized by *Pseudomonas*
A:Reference number: S69336, MUID:95262702, PMID:7744061
A:Accession: S69336

A:Molecule type: DNA
A:Residues: 1-533 <BEI>
A:Cross-references: EMBL:248540; NID:g2440146; PIDN:CAA88421.1; PID:g6656884
A:Experimental source: strain PA01
A:Accession: 560350

A: Molecule type: protein
A: Residues: 2-27; 95-101; 131-155; 159-169; 188-195; 245-263; 287-302; 340-350; 383-388; 407-419
R: Dietrichs, T.; Miesch, C.; Hummer, Johann, J.; Schmidt, B.; Kertesz, M.A.; von Figura, K.
J. Biol. Chem. 273, 25560-25564, 1998
Title: Posttranslational formation of formyloligocysteine in prokaryotic sulfatases by modification

A:Reference number: A59074; MUID:98421466; PMID:9748219
A:Contents: annotation, post-translational modification
A:Note: confirmation of 3-oxoalanine active site, referred to as formylglycine
C:Genetics:
C:Gene: at5A

C;Superfamily: animal sulfatase
C;Keywords: sulfuric ester hydrolase
F;51/Modified site: 3-oxoalanine (Cys) #status experimental

Query Match	21.9%	Score 57	2	Length 533
Best Local Similarity	36.1%	Pred. No.	24	
Matches	13	Conservative	8	Mismatches 13
				Indels 2
				Gaps 1

```
QY      13 LQRPHLMDGGMRHFSAGPELLRQDKPRSGSTG 48  
        |||::: | : |||||::: |  
Db      214 LQAPREYE--KRYGRYDAGPEALRQERLARLKELG 247
```

RESULT 13
D83622
arylsulfatase PA0183 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83622

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.; Bader, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: D83622

A: Molecule type: DNA
A: Residues: 1-536 <STO>
A: Cross-references: GB:AE004456; GB:AE004091; NID:G9946013; PIDN:AGC03573.1; GSPDB:GM001202401
A: Experimental source: strain PA01

A;Gene: atsa; PA0183
C;Superfamily: animal sulfatase

Query Match	21.98;	Score	57;	DB	2;	Length	536;
Best Local Similarity	36.18;	Pred. No.	24;				
Matches	13;	Conservative	8;	Mismatches	13;	Indels	2;
						Gaps	1;

```
QY      13 LÖRPÖHLMDÖGÖMRHSFSAPELLRÖDKRPRSGSTG 48
        ||:::::|::|::|::|::|::|::|
Db      214 LÖAPREIVE--KYRGRYDAGPEALRÖERLARLEKELG 247
```

RESULT 14
YXHUT

C:Species: Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence J01120:1-13342
C:Accession: A23047: 155318: A223299: A3342
R:Yakelshl, K.; Kaneda, S.; Ayusawa, D.; Shimizu, K.; Gotoh, O.; Seno, T.
C:Institute: synchrn (bc 2.1.1.43) Human

A:Title: Nucleotide sequence of a functional cDNA for human thymidylate synthase
A:Reference number: A23047; M0ID:85215597; PMID:2987839
A:Accession: A23047
A:Molecule type: mRNA
A:Position: 1-313 (nt)

A: Cross-references: EMBL: X02308; NID: G37478; PIDD: CA26178.1; PID: G37479
R: Kaneda, S.; Nalbantoglu, J.; Taketishi, K.; Shimizu, K.; Gotoh, O.; Seno, T.; Ayus
J. Biol. Chem. 265, 20277-20284, 1990
A: Title: Structural and Functional Analysis of the Human Thymidylate Synthase Gene
Reference number: 155318; MIM: 61056070; DMT: 2243092

A:Accession: I55318
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-313 <RES>
A:Cross-references: NITD:g220135: PTND:BA000472.1: PTD:g220136

R:Takeichi, K.; Ikemeda, S.; Aiyusawa, D.; Shimizu, K.; Gotoh, O.; Seno, T.
J. Biochem. 106, 575-583, 1989
A:Title: Human Ethymidylate synthase gene: Isolation of phage clones which cover a fun
A:Reference number: J00120; MUID:90110051; PMID:2532645
A:Accession: J00120

A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-68 <TA2>
A:Cross-references: GB:D00517, NID:9220133, PIDN:BA00404.1, PID:92160415
A:Shimizu, K.; Ayusawa, D.; Takeishi, K.; Sano, T.

U. Biochem. 91/1845-850, 1985
A:Title: Purification and NH2-terminal amino acid sequence of human thymidylate synthase
A:Reference number: A22393; MUID:85261174; PMID:3839505
A:Accession: A22393
A:Molecule type: protein

A:Residues: 7-23 $\text{C}_{12}\text{H}_{23}\text{NO}_2$
R:Davidsson, V.J.; Sitawaraporn, W.; Santl, D.V.
J. Biol. Chem. 264, 9145-9148, 1989
A:Title: Expression of human thymidylate synthase in *Escherichia coli*
R:Reference number: A33842; MOID:89255401; PMID:2656695

A:Accession: M35042
A:Molecule type: protein
A:Residues: 2-10 <DAV>
C:Genetics:
A:Gene: GDB:TYMS

C:Keywords: deoxyribonucleotide biosynthesis; methyltransferase
A:Superfamily: thymidylate synthase; thymidylate synthase homology
A:introns: 69/1; 93/3; 152/1; 186/1; 244/3; 268/3
A:Map position: 18p11.32-18p11.32
A:Cross references: SDBP,12403; OMIM,100000

Query Match	21.7%	Score 56.5;	DB 1;	Length 313;
Best Local Similarity	34.0%	Pred. NO. 15;		

QY 2 PLGRGDSPELLORPH--LMDOGMRHSFAGPELLRODKRPSGS 46
| | | : | | | : | | | : | | | :
Db 12 PLPRADGRDAERPRHGELOYLGQONHLRCG---VRDDKRGSTGT 55

RESULT 15
D86350
F8K7.13 protein - Arabidopsis thaliana

RESULT 15
D86350
F8K7.13 protein - Arabidopsis thaliana


```
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C:Accession: DB63350
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chlu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: DB6350
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-807 <STO>
A:Cross-references: GB:AE005172; NID:g5263321; PIDN:AAD41423.1; GSPOB:GN00141
Genetics:
A:Map position: 1

Query Match      21.7% Score 56.5; DB 2; Length 807;
Best Local Similarity 29.6%; Pred. No. 44;
Matches 16; Conservative 7; Mismatches 22; Indels 9; Gaps 1;

Oy      5 PRGDSPLTLPORPILM-----DOGOMHSFSAGPELLROKRPKRSSTGS 49
          |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      708 PGESNNRYOVAKMFMAHQDQDQDQDQDQDQSFGRRLPLMIQTNAASGTAS 761
```

Search completed: June 17, 2003, 12:01:16
Job time : 7.86764 secs

THIS PAGE BLANK (00070)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:41; Search time 2.94143 Seconds
(Without alignments)
690.935 Million cell updates/sec

Title: US-09-991-681-31
Perfect score: 260
Sequence: 1 EPLGPRGQDSPLLRPOHLM.....SAGPELLRQDKRRPSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	23.5	250	1 HXB9_HUMAN	P17482 homo sapien
2	59.5	22.9	2004	1 MOZ_HUMAN	Q92794 homo sapien
3	58	22.7	271	1 HMPH_MOUSE	P43120 mus musculu
4	58	22.3	270	1 HMPH_HUMAN	Q03014 homo sapien
5	58	22.3	943	1 NFCL_HUMAN	Q05644 homo sapien
6	57	21.9	485	1 ONC2_HUMAN	Q05948 homo sapien
7	57	21.9	535	1 ARS_PSEAE	P51691 pseudomonas
8	56.5	21.7	312	1 TTSY_HUMAN	P04818 homo sapien
9	56	21.5	163	1 PINI_HUMAN	Q13526 homo sapien
10	56	21.5	277	1 HMPH_CHICK	P37418 salmoneilla
11	55.5	21.3	455	1 MEHE_SALTY	Q05502 gallus galli
12	55.5	21.3	622	1 ABS_HUMAN	Q09199 homo sapien
13	55	21.2	186	1 RL18_YEAST	P07279 saccharomyc
14	55	21.2	1629	1 ARS9_HUMAN	Q09244 homo sapien
15	54.5	21.0	170	1 ZSS2_MOUSE	P15458 arabidopsis
16	54	20.8	165	1 PINI_MOUSE	Q09477 mus musculu
17	54	20.8	473	1 ULL3_HOYA	P16755 human cytom
18	54	20.8	623	1 PRT1_DROME	P51022 drosophila
19	54	20.8	2212	1 CCAA_RAT	P54282 rattus norv
20	53.5	20.6	1675	1 CLH1_HUMAN	Q00610 homo sapien
21	53.5	20.6	1675	1 CLH1_BOVIN	P49951 bos taurus
22	53.5	20.6	1675	1 CLH1_RAT	P11442 rattus norv
23	53.5	20.6	3063	1 POLG_PVN	P18247 p genome po
24	53	20.4	677	1 AGL3_ARATH	Q04777 arabidopsis
25	53	20.4	677	1 RGS9_RAT	P49805 rattus norv
26	53	20.4	1336	1 W146_HUMAN	Q06038 homo sapien
27	52.5	20.2	723	1 ICAL_SHEEP	Q09208 ovis aries
28	52.5	20.2	856	1 POLG_PYO	P22602 potato viru
29	52.5	20.2	3301	1 CLR3_MOUSE	Q03885 mus musculu
30	52	20.0	262	1 ASH2_MOUSE	Q03885 mus musculu
31	52	20.0	334	1 PYRB_PSEAE	Q05653 pseudomonas
32	52	20.0	334	1 PYRB_PSEAE	Q05711 pseudomonas
33	52	20.0	675	1 RGS9_MOUSE	Q04828 mus musculu

34	52	20.0	704	1 SYN1_RAT	P09951 rattus norv
35	52	20.0	706	1 SYN1_BOVIN	P17599 bos taurus
36	52	20.0	809	1 PPAP_XYLEA	Q09644 xylella fas
37	52	20.0	1640	1 CHH2_HUMAN	P53675 homo sapien
38	51.5	19.8	164	1 ZSS3_ARATH	P15459 arabidopsis
39	51.5	19.8	275	1 VINT_FRG3V	P29164 frog virus
40	51.5	19.8	759	1 MAO2_ECOLI	P76558 escherichia
41	51.5	19.8	2832	1 NOV8_RHIME	P20471 rhizobium m
42	51	19.6	346	1 MGB4_HUMAN	Q15481 homo sapien
43	51	19.6	501	1 CYG2_HUMAN	Q15488 homo sapien
44	51	19.6	550	1 YW40_YEAST	Q03212 saccharomyc
45	51	19.6	839	1 VPPI_MOUSE	Q02194 mus musculu

ALIGNMENTS

RESULT 1
HXB9_HUMAN STANDARD: PRT: 250 AA.
AC P17482: 09H11;
DT 01-AUG-1990 (Rel. 15, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B9 (Hox-2E) (Hox-2.5).
GN HOXB9 OR HOX2E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kosaki K., Kosaki R., Suzuki T., Yoshinashi H., Sasaki K., Matsuo N.;
RT "A complete mutation analysis panel of human HOX genes."
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL TISSUE=Muscle;
RN [3]
RP Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 173-250 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89378558; PubMed=2570724;
RA Giampolo A., Acampora D., Zappavigna V., Pannese M.,
RA D'Esposito M., Care A., Falletta A., Stornaiuolo A., Russo G.,
RA Simeone A., Boncinelli E., Peschle C.;
RT "Differential expression of human HOX-2 genes along the anterior-posterior axis in embryonic central nervous system."
RL Differentiation 40:191-197(1989).
RN [5]
RP SEQUENCE OF 185-250 FROM N.A.
RX MEDLINE=90215256; PubMed=2576652;
RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Caffiero M., Falletta A., Simeone A.;
RT "Organization of human class I homeobox genes."
RL Genome 31:745-756(1989).
RN [6]
RP FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN WHOLE EMBRYOS AND FETUSES AT
CC 5-9 WEEKS FROM CONCEPTION.
CC -1- SIMILARITY: BELONGS TO THE ARD-B HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

DR InterPro: IPR002717; MOZ_SAS.
DR InterPro: IPR001965; ZnF_PHD.
DR Pfam: PF006528; PHD_2.
DR Pfam: PF01853; MOZ_SAS_1.
DR SMART: SM00526; H1S_1.
DR SMART: SM00249; PHD_2.
DR PROSITE: PS01359; ZF_PHD_1; 1.
DR PROSITE: PS50016; ZF_PHD_2; 2.
KW Proto-oncogene; Chromosomal translocation; Zinc-finger; Repeat;.
KM Nuclear protein.
FT ZN_FING 206 265 PHD-TYPE 1.
FT ZN_FING 259 313 PHD-TYPE 2.
FT FT DOMAIN 371 379 POLY-SER.
FT ZN_FING 538 560 POLY-SER.
FT DOMAIN 788 801 POLY-GLU.
FT DOMAIN 989 995 POLY-GLU.
FT DOMAIN 1019 1026 POLY-ARG.
FT DOMAIN 1069 1078 POLY-GLU.
FT DOMAIN 1147 1150 POLY-LYS.
FT DOMAIN 1221 1242 GLU-RICH.
FT DOMAIN 1267 1302 GLU-RICH.
FT DOMAIN 1411 1414 POLY-GLU.
FT DOMAIN 1593 1597 POLY-SER.
FT DOMAIN 1643 1704 GLN/PRO-RICH.
FT DOMAIN 1897 1977 MET-RICH.
FT SITE 1546 1547 BREAKPOINT FOR TRANSLLOCATION TO FORM MOZ-CBP.
SQ SEQUENCE 2004 AA; 225054 MW; 9FFBAC3792854BA CRC64;

Query Match 22.9%; Score 59.5; DB 1; Length 2004;
Best Local Similarity 32.1%; Pred. No. 36;
Matches 17; Conservative 9; Mismatches 16; Indels 11; Gaps 2;

QY 5 PRGDSPLLRP-----QHLMDDQGMRHSFSGPELLRDKRPKSGTGS 49
   1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1679 PQQQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQP 1728
                                     -EIPESGSTGN

RESULT 3
HMPH_MOUSE STANDARD; PRT; 271 AA.
ID HMPH_MOUSE
AC P43120;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein PRH (Hematopoietically expressed homeobox) (Homeobox protein Hex).
DE HHX OR PRHX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathii; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=93219088; Pubmed=8096636;
RA Bedford F.K., Ashworth A., Enver T., Wiedemann L.M.;
RT "Hex: a novel homeobox gene expressed during haematopoiesis and conserved between mouse and human.";
RL Nucleic Acids Res. 21:1245-1249(1993).
CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE IN HEMATOPOIETIC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING HEMATOPOIESIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; 221524; CAA79729.1; -
 DR HSSP; P22808; INK3.
 DR TRANSFAC; T03417; -
 DR MGD; MGI:96086; Hhex.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 DR Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 1 134 PRO-RICH.
 FT DNA_BIND 138 197 HOMEBOX.
 SO SEQUENCE 271 AA; 29986 MM; 37676557A71C962 CRC64;
 Query Match 22.7%; Score 59; DB 1; Length 271;
 Best Local Similarity 45.7%; Pred. No. 4.1;
 Matches 16; Conservative 2; Mismatches 17; Indels 0; Gaps 0;
 QY 1 EPLGPRGDSPLRQPHLMDQGMRHSFSAGPEL 35
 Db 119 DPLGKPLMSPLRQPHLKRKGQVRFSDQTVEL 153
 RESULT 4
 HMPH_HUMAN STANDARD; PRT; 270 AA.
 AC 003014;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUN-1994 (Rel. 29, Last sequence update)
 DE Homeobox protein PRH (Hematopoietically expressed homeobox) (Homeobox
 DE protein HEM).
 GN HHEX OR PRHX OR PRH OR HEX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=93087175; PubMed=1360645;
 RA Crompton M.R., Bartlett T.J., Macgregor A.D., Manfioletti G.,
 RA Burtati E., Giancotti V., Goodwin G.H.;
 RT "Identification of a novel vertebrate homeobox gene expressed in
 RT haematopoietic cells.";
 RL Nucleic Acids Res. 20:5661-5667(1992).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93384629; PubMed=8103988;
 RA Hromas R.A., Collins S.J., Radich J.;
 RT "PCR cloning of an orphan homeobox gene (PRH) preferentially
 RT expressed in myeloid and liver cells.";
 RL Biochem. Biophys. Res. Commun. 195:976-983(1993).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA Strausberg R.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE OF 122-270 FROM N.A.
 RX MEDLINE=93219088; PubMed=8096636;
 RA Bedford F.K., Ashworth A., Enver T., Wiedemann L.M.;
 RT "Hx: a novel homeobox gene expressed during haematopoiesis and
 RT conserved between mouse and human.";
 RL Nucleic Acids Res. 21:1245-1249(1993).
 RN [5]
 RN 3D-STRUCTURE MODELING OF 136-196.
 RX MEDLINE=94259190; PubMed=7911091;
 RA Neidie S., Goodwin G.H.;
 RT "A homology-based molecular model of the proline-rich homeodomain
 RT protein Prh, from haematopoietic cells.";
 RL FEBS Lett. 345:93-98(1994).

CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
 CC IN HEMATOPOIETIC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: LIVER AND PROMYELOCYTIC LEUKEMIA CELL
 CC LINE HL60.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING HEMATOPOIESIS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X67235; CAA47661.1; -
 CC EMBL; L16499; AAA02988.1; -
 CC EMBL; BC015110; AAH15110.1; -
 CC EMBL; 221533; CAA79730.1; -
 CC PIR; S26799; S26799.
 CC HSSP; P22808; INK3.
 CC TRANSFAC; T02092; -
 CC Genew; HGNC:4901; HHEX.
 CC MIM; 604420; -
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS50071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 1 133 PRO-RICH.
 FT DNA_BIND 137 196 HOMEBOX.
 FT DNA_BIND 115 115 L->V (IN REF. 2).
 FT SEQUENCE 270 AA; 30021 MM; 9C16B6D494475FC CRC64;
 Query Match 22.3%; Score 58; DB 1; Length 270;
 Best Local Similarity 50.0%; Pred. No. 5.4;
 Matches 14; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
 QY 1 EPLGPRGDSPLRQPHLMDQGMRHS 28
 Db 118 DPLGKPLMSPLRQPHLKRKGQVRFSDQTVEL 145
 RESULT 5
 NFCL_HUMAN STANDARD; PRT; 943 AA.
 ID NFCL_HUMAN
 AC 095644; Q12865; Q15793;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nuclear factor of activated T-cells, cytoplasmic 1 (NFAT transcription
 DE complex cytosolic component) (NF-ATc1) (NF-ATc).
 GN NFATC1 OR NFATC OR NFAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RX SEQUENCE FROM N.A. (ISOFORM A-ALPHA).
 RC TISSUE=T-cell, and Peripheral blood lymphocytes;
 RX MEDLINE=94261186; PubMed=8202141;
 RA Northrop J.P., Ho S.N., Chen L., Thomas D.J., Timmerman L.A.,
 RA Nolan G.P., Admon A., Crabtree G.R.;
 RT "NF-AT components define a family of transcription factors targeted in
 RT T-cell activation.";
 RL Nature 369:497-502(1994).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM B-BETA).
 RC TISSUE=B-cell;
 RX MEDLINE=96355439; PubMed=8702849;

RA Park J., Takeuchi A., Sharma S. :
RA "Characterization of a new isoform of the NFAT (nuclear factor of
RT activated T cells) gene family member NFATC.,"
RL J. Biol. Chem. 271:20914-20921(1996).
RN [3]
RN
RP ERRATUM.
RA Park J., Takeuchi A., Sharma S. :
RL J. Biol. Chem. 271:33705-33705(1996).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A-ALPHA; B-ALPHA AND C-BETA).
RC TISSUE-B-cell lymphoma.
RX MEDLINE-99170294; PubMed-10072078;
RA Chuvpilo S., Zlummer M., Kerstan A., Gloeckner J., Avots A., Escher C.,
RA Fischer C., Jansekina I., Jankevics E., Berberich-Stebelt F.,
RA Schmitt E., Seifling E.;
RT "Alternative polyadenylation events contribute to the induction of
RT NF-ATc in effector T cells.",
RN Immunity 10:261-269(1999).
RN [5]
RN MUTAGENESIS.
RX MEDLINE-20119316; PubMed-10652349;
RA Porter C.M., Havens M.A., Clippstone N.A. ;
RT "Identification of amino acid residues and protein kinases involved in
RL the regulation of NFATc subcellular localization.",
RL J. Biol. Chem. 275:3543-3551(2000).
RN [6]
RN ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RX MEDLINE-99288090; PubMed-10358178;
RA Chuvpilo S., Avots A., Berberich-Stebelt F., Gloeckner J., Fischer C.,
RA Kerstan A., Escher C., Jansekina I., Hlupek F., Jankevics E.,
RA Brietz T., Seifling E.;
RT "Multiple NF-ATc isoforms with individual transcriptional properties
RT are synthesized in T lymphocytes.",
RL Immunol. 162:7294-7301(1999).
RN [7]
RN REVIEW.
RP MEDLINE-99189746; PubMed-10089876;
RX Crabtree G.R. ;
RA "Genetic signals and specific outcomes: signaling through Ca2+,
RT calcineurin, and NF-AT.",
RL Cell 96:611-614(1999).
RN [8]
RN CELL FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 OR IL-4
CC GENE TRANSCRIPTION. ALSO CONTROL GENE EXPRESSION IN EMBRYONIC
CC CARDIAC CELLS. COULD REGULATE NOT ONLY THE ACTIVATION AND
CC PROLIFERATION BUT ALSO THE DIFFERENTIATION AND PROGRAMMED DEATH OF
CC T-LYMPHOCYTES AS WELL AS LYMPHOID AND NONLYMPHOID CELLS.
CC
CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC THAT CONSISTS OF AT LEAST TWO COMPONENTS. A PRE-EXISTING
CC CYTOSOLIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
CC ACTIVATING PROTEIN-1 FAMILY. MAP, GATA4 AND CBP/100 CAN ALSO BIND
CC THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.
CC
CC -1- SUBCELLULAR LOCATION: CYTOSOLIC FOR THE PHOSPHORYLATED FORM AND
CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS; A-ALPHA, A-BETA, B-
CC ALPHA, B-BETA, C-ALPHA (SHOWN HERE) AND C-BETA. ARE PRODUCED BY
CC ALTERNATIVE SPLICING. THERE MIGHT BE ADDITIONAL ISOFORMS PRODUCED
CC BY ALTERNATIVE INITIATION AT MET-37 OF ISOFORMS ALPHA. NFATC/C IS
CC THE STRONGEST ACTIVATOR OF GENE TRANSCRIPTION, FOLLOWED BY
CC NFATC/A, WHEREAS NFATC/B IS THE WEAKEST. NFATC/B AND NFATC/C, BOTH
CC PRESENT IN T-CELLS CAN MODULATE THEIR TRANSCRIPTIONAL ACTIVITY.
CC
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS, PERIPHERAL LEUKOCYTES AS
CC T-CELLS AND SPLEEN. NFATC/A IS PREFERENTIALLY EXPRESSED IN
CC EFFECTOR T-CELLS (THYMUS AND PERIPHERAL LEUKOCYTES) WHEREAS
CC NFATC/B AND NFATC/C ARE PREFERENTIALLY EXPRESSED IN NAIVE T-CELLS
CC (SPLEEN). NFATC/B IS EXPRESSED IN NAIVE T-CELLS AFTER FIRST
CC ANTIGEN EXPOSURE AND ISOFORM A IS EXPRESSED IN EFFECTOR T-CELLS
CC AFTER SECOND ANTIGEN EXPOSURE.

CC	-1- INDUCTION: ONLY NEAT/C/A IS INDECIPIRALLY EXPRESSED IN T LYMPHOCYTES UPON ACTIVATION OF THE T-CELL RECEPTOR (TCR) COMPLEX. INDUCED AFTER CO-ADDITION OF PHORBOL 12-MIRISTATE 13-ACETATE (PMA) AND IONOMYCIN. ALSO INDUCED AFTER CO-ADDITION OF 12-O-TETRADECANOYLPHORBOL-13-ACETATE (TPA) AND IONOMYCIN. WEAKLY INDUCED WITH PMA, IONOMYCIN AND CYCLOSOPHRIN A.
CC	-1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND COOPERATIVE INTERACTIONS WITH API FACTORS.
CC	-1- DOMAIN: THE N-TERMINAL TRANSACTIVATION DOMAIN (TAD-A) BINDS TO AND IS ACTIVATED BY CEB/P300. THE DEPHOSPHORYLATED FORM CONTAINS TWO UNMARKED NUCLEAR LOCALIZATION SIGNALS (NLS), WHICH ALLOW TRANSLLOCATION OF THE PROTEIN TO THE NUCLEUS.
CC	-1- DOMAIN: NEAT/C HAS A C-TERMINAL PART WITH AN ADDITIONAL TRANS-ACTIVATOR. NEAT/B HAS A SHORTER C-TERMINAL PART WITHOUT COMPLETE TAD-B WHICH ACTS AS A TRANSCRIPTIONAL REPRESSOR.
CC	-1- PTM: PHOSPHORYLATED BY NEAT-C-KINASE; DEPHOSPHORYLATED BY CALCINEURIN.
CC	-1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC	-----
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION between the Swiss Institute of Bioinformatics and the EMBL OUTSTATION - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/announce/ or send an email to license@isdb.slb.ch).
CC	-----
DR	EMBL; U08015; AAA19601.1; -;
DR	EMBL; U59736; AAC50869.1; -;
DR	EMBL; U80917; AAD00450.1; -;
DR	EMBL; U80918; AAD00451.1; -;
DR	EMBL; U80919; AAD00452.1; -;
DR	TRANSFAC; T01945; -;
DR	GeneW; HGNC:7775; NEATC1.
DR	MIM; 600489; -;
DR	InterPro; IPRO02909; IPT_TIG.
DR	InterPro; IPRO00451; NF_Rel_dor_fam.
DR	Pfam; PF00534; RHD; 1.
DR	Pfam; PF01833; TIG; 1.
DR	SMART; SMO0429; IPT; 1.
DR	PROSITE; PS01204; REL_1; FALSE_NEG.
KW	PROSITE; PSS0254; REL_2; 1.
KW	Transcription regulation; Activator; Repressor; Nuclear protein; DNA-binding; Alternative splicing; Phosphorylation; Repeat.
FT	DNAIN 118 123
FT	DNAIN 126 218
FT	DNAIN 203 298
FT	REPEAT 203 219
FT	REPEAT 233 249
FT	REPEAT 282 298
FT	DNAIN 265 267
FT	DNAIN 310 321
FT	DNAIN 439 446
FT	DNAIN 682 684
FT	DNAIN 703 943
FT	DNAIN 924 933
FT	MOD_RES 117 117
FT	VARSPLIC 1 42
FT	----
FT	VARSPLIC 698 716
FT	----
FT	* VARSPLIC 717 943
FT	VARSPIC 826 943
FT	MUTAGEN 169 169
FT	MUTAGEN 172 172
FT	MUTAGEN 187 187

FT FT LOCALIZATION.
 FT CONFLICT 232 232 G -> S (IN REF. 2).
 FT CONFLICT 235 235 R -> O (IN REF. 1).
 SO SEQUENCE 943 AA; 101242 MW; E72F8B10ECB2D66 CRC64;
 Query Match 22.3%; Score 58; DB 1; Length 943;
 Best Local Similarity 34.5%; Pred. No. 23;
 Matches 19; Conservative 4; Mismatches 18; Indels 14; Gaps 2;

OY 2 PLGRGDSPLLQ-----RQHLMDGOMRHSFSGPELL---RQDKRP 42
 DB 850 PLRPTGEPCTCLQPCSPACRPATGRPHLPSTVRDESPTAPRLPEVHEDGSP 904

RESULT 6
 ONC2_HUMAN STANDARD; PRT; 485 AA.
 ID ONC2_HUMAN
 AC 095948;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE One cut domain family member 2 (ONECUT-2 transcription factor) (OC-2).
 GN ONECUT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9915796;
 RA Jacquemin P., Lannoy V., Rousseau G.G., Lemaigre F.P.;
 "OC-2, a novel mammalian member of the ONECUT class of homeobox
 transcription factors whose function in liver partially overlaps with
 that of hepatocyte nuclear factor-6.";
 RL J. Biol. Chem. 274:2665-2671(1999).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. ACTIVATES THE TRANSCRIPTION
 CC OF A NUMBER OF LIVER GENES SUCH AS HNF3B.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE CUT HOMEBOX FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Y18198; CAB38253.1; -;
 DR TRANSFAC; T03259; -;
 DR Genew; HGNC:8139; ONECUT2.
 DR MIM; 604894; -;
 DR InterPro; IPR003350; Hmoec-CUT.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox.1.
 DR Pfam; PF02376; CUT.1.
 DR ProDom; PD000010; Homeobox.1.
 DR SMART; SM00389; HOX.1.
 DR PROSITE; PS00027; HOMEBOX_1; FALSE_NEG.
 DR PROSITE; PS00711; HOMEBOX_2; 1.
 KW Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
 KW Activator.
 FT DN_BIND 305 391 CUT.
 FT DN_BIND 407 466 HOMEBOX.
 FT DOMAIN 18 37 POLY-GLY.
 FT DOMAIN 62 66 POLY-PRO.
 FT DOMAIN 75 82 POLY-ALA.
 FT DOMAIN 152 165 POLY-HIS.
 FT DOMAIN 298 303 POLY-SER.
 SO SEQUENCE 485 AA; 52482 MW; AF21E052EFBE5DA1 CRC64;
 Query Match 21.9%; Score 57; DB 1; Length 485;

Best Local Similarity 36.4%; Pred. No. 14;
 Matches 16; Conservative 4; Mismatches 24; Indels 0; Gaps 0;

OY 6 RGDSPLLRPHLMDGOMRHSFSGPELLRQKRRPSSGSGS 49
 DB 260 RGLCTPPAAMSHLNGLHHPGHTQSHGVLAPSRERPPSSSSGS 303

RESULT 7
 ARS_PSEAE STANDARD; PRT; 535 AA.
 ID ARS_PSEAE
 AC P51691;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Arylsulfatase (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
 GN ARSA OR PA0183.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=95262702; PubMed=7744061;
 RA Bell S., Kehrl H., James P., Staudemann W., Cook A.M., Leisinger T.,
 Kertesz M.A.;
 "Purification and characterization of the arylsulfatase synthesized
 RT by Pseudomonas aeruginosa PAO during growth in sulfate-free medium
 RT and cloning of the arylsulfatase gene (atsa).";
 RL Eur. J. Biochem. 229:385-394(1995).
 RN [2]
 RP REVISIONS.
 RA Kertesz M.A.;
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 Gader R.L., Goltzy L., Tolentino E., Westbrock-Webman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Lapidig K., Lam R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.;
 "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 RT opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 CC -1- FUNCTION: SHOWS MAXIMAL ACTIVITY AT 57 DEGREES CELSIUS AND PH 8.9.
 CC INCUBATION THAT OVERSTEPS 20 MINUTES ABOVE 50 DEGREES CELSIUS
 CC LEADS TO ENZYME INACTIVATION.
 CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2O) -> a phenol + sulfate.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z48540; CAAB8421.2; -;
 DR EMBL; AE004456; AACG3573.1; -;
 DR HSSP; P15848; 1FSU.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase.1.
 DR PROSITE; PS00523; SULFATASE_1; 1.
 DR PROSITE; PS00149; SULFATASE_2; 1.
 KW Hydrolase; Complete proteome.
 FT INIT_MET 0 0

RT "A human peptidyl-prolyl isomerase essential for regulation of
 RT mitosis."
 RL Nature 380:544-547(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (FEb-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
 RX MEDLINE=97344079; PubMed=9200606;
 RA Ranganathan R., Lu K.P., Hunter T., Noel J.P.;
 RT "Structural and functional analysis of the mitotic rotamase Pin1
 RT suggests substrate recognition is phosphorylation dependent.";
 RL Cell 89:875-886(1997).
 CC -1- FUNCTION: Essential PPIase that regulates mitosis presumably by
 CC interacting with NIMA and attenuating its mitosis-promoting
 CC activity. Displays a preference for an acidic residue N-terminal
 CC to the isomerized proline bond. Catalyzing pSer/Thr-Pro cis/trans
 CC isomerizations.
 CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega-180) = peptidylproline
 CC (omega-0).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
 CC -1- SIMILARITY: CONTAINS 1 MW DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U49070; AAC50492.1; -;
 DR EMBL: BC002899; AAH02899.1; -;
 DR PDB: 1PIN; 25-NOV-98.
 DR Genew: HGNC:8988; PIN1.
 DR MIM: 601052; -;
 DR InterPro: IPR000297; Rotamase.
 DR InterPro: IPR002349; MW_Rsp5_MWP.
 DR InterPro: IPR001202; MW_Rsp5_MWP.
 DR Pfam: PF00397; MW; 1.
 DR Pfam: PF00639; Rotamase; 1.
 DR PRINTS: PR00403; MWDOMAIN.
 DR SMART: SM00456; MW; 1.
 DR PROSITE: PS01096; PPIC_PPIASE_1; 1.
 DR PROSITE: PS01198; PPIC_PPIASE_2; 1.
 DR PROSITE: PS01159; MW_DOMAIN_1; 1.
 DR PROSITE: PS50020; MW_DOMAIN_2; 1.
 KM Isomerase; Rotamase; Nuclear protein; Cell cycle; 3D-structure.
 FT DOMAIN 5 39 MW
 FT DOMAIN 52 163 PPIC.
 SO SEQUENCE 163 AA; 18243 MW; 35391AF40B7D1E13 CRC64;
 Query Match 21.5%; Score 56; DB 1; Length 163;
 Best Local Similarity 36.6%; Pred. No. 5.2;
 Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;
 QY 4 GPRGDSPLLRPQHMDQGMRSFSAPELLRQDKRPRS 44
 DB 45 GKNGQEPARYRCSHL-----VKHSQSRPSSMROEKITRT 81
 RESULT 10
 HMPH_CHICK STANDARD; PRT; 277 AA.
 ID HMPH_CHICK
 AC O05502;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Homeobox protein PRH.
 OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CC OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93087175; PubMed=1360645;
 RA Crompton M.R., Bartlett T.J., Macgregor A.D., Manfioletti G.,
 RA Buratti E., Giannotti V., Goodwin G.H.;
 RT "Identification of a novel vertebrate homeobox gene expressed in
 RT hematopoietic cells.";
 RL Nucleic Acids Res. 20:5661-5667(1992).
 CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
 CC IN HEMATOPOIETIC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: IN ALL HEMATOPOIETIC TISSUES EXCEPT
 CC PERIPHERAL BLOOD ERYTHROCYTES AND IN THE LIVER AND LUNG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64711; CAA45966.1; -;
 DR HSRP: P22808; IMK3.
 DR TRANSFAC: T02091; -;
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR PRODOM: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS50071; HOMEBOX_2; 1.
 KM Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 1 140
 FT DOMAIN 144 203 HOMEBOX.
 SO SEQUENCE 277 AA; 30213 MW; BE744C1A3FB9F9FC CRC64;
 Query Match 21.5%; Score 56; DB 1; Length 277;
 Best Local Similarity 46.4%; Pred. No. 9.7;
 Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 QY 1 EPISRGDSPLLRPQHMDQGMRS 28
 DB 125 DPILKPLWSPFIQRPILKRGQVRS 152
 RESULT 11
 MENE_SALTY STANDARD; PRT; 455 AA.
 ID MENE_SALTY
 AC P37418;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26) (OSB-CoA synthetase)
 DE (O-succinyl)benzoate-CoA synthase).
 GN MENE OR STM2305.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Salmonella.
 CC OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portolillo S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;

```

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT L2."
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 228-455 FROM N.A.
RC STRAIN-LT2;
RX MEDLINE-94266712; Pubmed-8206837;
RA Roland K.L., Esther C.R., Spitznagel J.K.;
RT "Isolation and characterization of a gene, pmwD, from Salmonella
RT Typhimurium that confers resistance to polymyxin when expressed in
RT multiple copies."
RL J. Bacteriol. 176:3589-3597(1994).
CC -1- FUNCTION: O-SUCCINYL-BENZOIC ACID (OSB) TO O-SUCCINYL-BENZOYL-COA
CC (OSB-COA).
CC -1- CATALYTIC ACTIVITY: ATP + O-succinylbenzoate + CoA = AMP +
CC diphosphate + O-succinylbenzoyl-CoA.
CC -1- PATHWAY: Menquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008803; AAL21206.1; -
DR EMBL: U02281; AAA21323.1; -
DR StyGene: SG10221; menE.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 2.
DR PROSITE: PS00455; AMP-BINDING; 1.
DR Menquinone biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 455 AA; 50244 MW; BC830AB7BFA0F42 CRC64;

Query Match 21.3%; Score 55.5; DB 1; Length 455;
Best Local Similarity 40.5%; Pred. No. 20;
Matches 17; Conservative 4; Mismatches 18; Indels 3; Gaps 2;

QY 2 PLGPGQDSPLLORPQ-HLMDQGMRSFSGAPPELLRODKRP 42
DB 171 PFGP--QDDWLISLPFHVSGGIMRWLPRGAMRYRDKOP 210

RESULT 12
ABS_HUMAN STANDARD: PRT; 622 AA.
AC 09UJY9; Q96BK6; Q9NW04; Q96K05;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE DEAD-box protein abstract homolog.
GN ABS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20076860; Pubmed-10607561;
RA Ition U., Lepin M.;
RT "Developmental and cell biological functions of the Drosophila DEAD-
RT box protein abstract."
RL Curr. Biol. 9:1373-1381(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Tisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Magatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,

```

```

RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakatsuki A., Nakamura Y., Nagahari K., Masuno Y.,
RA Nimoto K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Skin;
RA Strausberg R.;
RT Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE ATP-DEPENDENT RNA HELICASE. IS REQUIRED DURING
CC POST-TRANSCRIPTIONAL GENE EXPRESSION.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF195417; AAR04150.1; -
DR EMBL: AK001255; BAA91585.1; -
DR EMBL: AK027768; BAB55355.1; -
DR EMBL: BC015476; AAI15476.1; -
DR HSSP: Q58083; 1HV8.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR001878; ZnF_CCHC.
DR Pfam: PF00098; zf-CCHC; 1.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR SMART: SM00343; ZnF_C2HC; 1.
DR PROSITE: PS00039; DEAD-ATP-HELICASE; FALSE_NEG.
DR ATP-binding; RNA-binding; Helicase; Nuclear protein.
FT NP_BIND 225 232
FT SITE 344 347
FT SITE 344 347
FT CONFLICT 17 43
FT CONFLICT 56 56
FT CONFLICT 64 64
FT CONFLICT 64 64
FT CONFLICT 165 165
FT CONFLICT 191 191
FT CONFLICT 352 352
FT CONFLICT 552 552
SQ SEQUENCE 622 AA; 69837 MW; E0A328724E0DF99A CRC64;

Query Match 21.3%; Score 55.5; DB 1; Length 622;
Best Local Similarity 31.9%; Pred. No. 28;
Matches 15; Conservative 10; Mismatches 21; Indels 1; Gaps 1;

QY 2 PLGPGQDSPLLORPQ-HLMDQGMRSFSGAPPELLRODKRPSGSG 48
DB 78 PLGPGQSVS-LLDQHQHLKEAKRSKSAKEKLEKEKILLESVAEG 123

RESULT 13
RL18_YEAST STANDARD: PRT; 186 AA.
ID RL18_YEAST
AC P07279;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L18 (Rp28).
GN (RPL18A OR RP28A OR YOL120C) AND (RPL18B OR RP28B OR YNL301C OR
GN N0425).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. (RPL18A AND RPL18B).
 RX MEDLINE=85037916; PubMed=6387623;
 RA Molenaar C.M.T., Mout R., Jansen A.E.M., Mager W.H.,
 RA Plant R.J., Donovan D.M., Pearson N.J.;
 RT "Structure and organization of two linked ribosomal protein genes in
 RT yeast.";
 RL Nucleic Acids Res. 12:7345-7358(1984).
 RN [2]
 RP SEQUENCE FROM N.A. (RPL18B).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96132033; PubMed=8553702;
 RA Maurer K.C.T., Urbanus J.H.M., Plant R.J.;
 RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV
 RT carrying a ribosomal protein gene cluster, the genes encoding a
 RT plasma membrane protein and a subunit of replication factor C, and a
 RT novel putative serine/threonine protein kinase gene.";
 RL Yeast 11:1303-1310(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (RPL18A).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97051591; PubMed=8896268;
 RA Lafuente M.J., Gamio F.J., Gancedo C.;
 RT "DNA sequence analysis of a 10 624 bp fragment of the left arm of
 RT chromosome XV from Saccharomyces cerevisiae reveals a RNA binding
 RT protein, a mitochondrial protein, two ribosomal proteins and two new
 RT open reading frames.";
 RL Yeast 12:1041-1045(1996).
 CC -1- SMISLARITY: BELONGS TO THE L18E FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- SMISLARITY: THERE ARE TWO GENES FOR L18 IN YEAST.
 CC -1- SMISLARITY: BELONGS TO THE L18E FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X01099; CAA25573.1; -;
 DR EMBL: X02635; CAA26481.1; -;
 DR EMBL: X01100; CAA25574.1; -;
 DR EMBL: U23084; AAC49097.1; -;
 DR EMBL: Z71577; CAA96219.1; -;
 DR EMBL: X95258; CAA64550.1; -;
 DR EMBL: Z74862; CAA99139.1; -;
 DR PIR: S05867; R5BY8E;
 DR SCD: S0005480; RPL18A.
 DR SCD: S0005245; RPL18B.
 DR InterPro: IPR000039; Ribosomal_L18e.
 DR Pfam: PF008828; Ribosomal_L18e; 1.
 DR PROSITE: PS01106; RIBOSOMAL_L18E; 1.
 KW Ribosomal protein; Multigene family.
 SQ SEQUENCE 186 AA; 20563 MW; D097B187F369EACD CRC64;
 Query Match 21.2%; Score 55; DB 1; Length 186;
 Best Local Similarity 36.4%; Pred. No. 8;
 Matches 16; Conservative 10; Mismatches 12; Indels 6; Gaps 3;
 OY 5 PRGDSPLQRPQHMDGGRHSFSAPELLRDKRPRSGSTG 48
 Db 132 PKQGNTLILGRPRN-SREAVRH-FGMGP---HKGAKPRILSTG 169
 RESULT 14
 AT93_HUMAN STANDARD; PRT; 1629 AA.
 AC 09P2N4; 09NR29;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).
 GN ADAMTS9 OR KIAA1312.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE=Fetal.
 RX MEDLINE=20396138; PubMed=10936055;
 RA Clark M.E., Kellner G.S., Turbeville L.A., Boyer A., Arden K.A.,
 RA Maki R.A.;
 RT "ADAMTS 9, a novel member of the ADAM-TS/metalloprotein gene
 RT family.";
 RL Genomics 67:343-350(2000).
 RN [2]
 RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (by similarity).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.
 CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,
 CC COLON. SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR
 CC THYMUS.
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
 CC SIMILARITY).
 CC -1- PFM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF261918; AAF89106.1; -;
 DR EMBL: AB037733; BAA92350.1; -;
 DR HSSP: P15167; IATL.
 DR MEROPS: M12_021; -;
 DR Genew: HGNC:13202; ADAMTS9.
 DR MIM: 605421; -;
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR002870; Pep_M12B_propep.
 DR InterPro: IPR001590; Repolysin.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF000090; tsp_1; 11.
 DR Pfam: PF01421; Repolysin; 1.
 DR Pfam: PF01562; Pep_M12B_propep; 1.
 DR SMART: SM00209; TSP1; 12.
 DR PROSITE: PS50215; ADAM_MEPRO; 1.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS50092; TSP1; 9.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolyase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Alternative splicing.

FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 287 BY SIMILARITY.
 FT CHAIN 288 1629 ADAMTS-9.
 FT DOMAIN 509 587 DISINTEGRIN-LIKE.
 FT DOMAIN 589 642 TSP TYPE-1.1.
 FT DOMAIN 645 752 CYS-RICH.
 FT DOMAIN 753 880 SPACER.
 FT DOMAIN 999 1053 TSP TYPE-1.2.
 FT DOMAIN 1056 1108 TSP TYPE-1.3.
 FT DOMAIN 1111 1156 TSP TYPE-1.4.
 FT DOMAIN 1184 1239 TSP TYPE-1.5.
 FT DOMAIN 1240 1295 TSP TYPE-1.6.
 FT DOMAIN 1332 1383 TSP TYPE-1.7.
 FT DOMAIN 1386 1439 TSP TYPE-1.8.
 FT DOMAIN 1445 1498 TSP TYPE-1.9.
 FT DOMAIN 1501 1554 TSP TYPE-1.10.
 FT DOMAIN 1562 1612 TSP TYPE-1.11.
 FT DOMAIN 88 96 POLY-SR.
 FT SITE 223 223 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 434 434 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 435 435 BY SIMILARITY.
 FT METAL 438 438 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 444 444 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 840 840 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPIC 1064 1072 CLYTCGCGH -> VRMGCGCFP (IN SHORT ISOFORM).
 FT VARSPIC 1073 1629 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 367 367 F -> L (IN REF. 1).
 SQ SEQUENCE 1629 AA; 182649 MW; C1C4CEFF58B8941F CRC64;

Query Match 21.28; Score 55; DB 1; Length 1629;
 Best Local Similarity 35.98; Pred. No. 1e+02;
 Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;

QY 8 QDSPLLRPQHLDQGMHSFAGPELLRQDRPRS GS 46
 11 : 11 : 11 : 11 : 11 : 11 :
 1288 QDCSMSPCQRTPTDGLAQHPQ-----NEDYRPRRAS 1320

RESULT 15
 2SS2_ARATH STANDARD; PRT; 170 AA.
 AC P15458;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 2S seed storage protein 2 precursor (2S albumin storage protein)
 GN (NMW02-2S albumin 2).
 DE AT252 OR AT4G27150 OR T24A18.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC Eucosids II: Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. C24;
 RA Kriebbers E., Hetdles L., de Clercq A., Seurinck J., Leemans J.,
 RA van Damme J., Segura M., Gheysen G., van Montagu M.,
 RA Vandekerckhove J.;
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. C24;
 RA Conceicao A.D.S., Kriebbers E.;

RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; Pubmed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Anstorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohneset U., Zimmermann M., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Rothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berneriser S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gelen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McLay K., Mayes R.,
 RA Pettett A., Rajadream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehert T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Grandtath K., Dauner D., Hertzl A.,
 RA Neumann S., Argitlou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Claudaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cheifor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bietke C.,
 RA Frishman D., Haese D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Hebermann K.,
 RA Parnell L., Dedia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cortes M., Abu-Threideh J.,
 RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pavin K., Hillier L.,
 RA Nelson J., Spleth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamat B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:769-777(1999).
 RN [4]
 RP SEQUENCE OF 1-90 AND 118-170 FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Raynal M., Grellet F., Landie M., Meyer Y., Cooke R., Delseny M.;
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M22034; AAA32744.1; -;
 DR EMBL: Z24745; CAA80871.1; -;
 DR EMBL: AL035680; CAB38845.1; -;
 DR EMBL: AL161566; CAB79570.1; -;

DR EMBL; Z17598; CAA79010.1; -;
DR EMBL; Z17594; CAA79008.1; -;
DR PIR; JAO162; NMMU2.
DR PIR; S34677; S34677.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000617; Napin.
DR Pfam; PF00234; tryd_alpha_amy1; 1.
DR PRINTS; PR00496; NAPIN.
DR ProDom; PD002498; Napin; 1.
DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Albumin; Signal; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 37
FT CHAIN 38 72 2S SEED STORAGE PROTEIN 2 SMALL SUBUNIT
FT PROPEP 73 88 (BY SIMILARITY).
FT CHAIN 89 170 2S SEED STORAGE PROTEIN 2 LARGE SUBUNIT
FT SEQUENCE 170 AA; 19361 MW; 0A0562D5FAB56275 CRC64;
OY 2 PLGPRGQDSPULQRPQHL-----MDGQMRHSFSAGPEL 35
DB 38 PMGPRQCKQKRFQOSQHLRACQKLRMQMRQMGGRGGPSL 76
Query Match 21.0%; Score 54.5; DB 1; Length 170;
Best Local Similarity 35.98; Pred. No. 8.3;
Matches 14; Conservative 5; Mismatches 15; Indels 5; Gaps 1;

Search completed: June 17, 2003, 11:56:49
Job time : 3.94143 secs

THIS PAGE BLANK (4870)

1	260	100.0	592	4	Q96CH9	Q96CH9 homo sapien
2	260	100.0	1807	4	Q9ULH6	Q9ULH6 homo sapien
3	63	24.2	462	16	Q9HUI9	Q9HUI9 pseudomonas
4	62.5	24.0	1655	5	Q24754	Q24754 drosophila
5	60	23.1	293	2	Q9X6T3	Q9X6T3 streptomyce
6	60	23.1	333	16	Q9RTI1	Q9RTI1 delnocioccus
7	59.5	22.9	345	17	Q9YBM1	Q9YBM1 aetopyrum p
8	59	22.7	271	11	Q9RIX2	Q9RIX2 mus musculuu
9	59	22.7	280	11	Q9ACT4	Q9ACT4 mus musculuu
10	59	22.7	381	16	Q9NCV1	Q9NCV1 mus musculuu
11	59	22.7	644	4	Q9NRK6	Q9ACT4 streptomyces
12	58.5	22.5	276	12	Q85375	Q9NRK6 homo sapien
13	58.5	22.5	580	3	Q9C2E4	Q85375 potato virtu
14	58.5	22.5	778	3	Q96WJ2	Q9C2E4 neurospora
15	58.5	22.5	3429	5	Q24553	Q96WJ2 kluyeromyc
16	58.5	22.5	3429	5	Q9W4Z5	Q24553 drosophila
						Q9W4Z5 drosophila

17	58	22.3	172	11	09CUI2	09cui2 sus muscula
18	58	22.3	221	6	095K08	095k08 sus scrofa
19	58	22.3	270	4	096CE9	096ce9 homo sapien
20	58	22.3	327	16	09A875	09a875 caulobacter
21	58	22.3	553	5	09N310	09n310 caenorhabdi
22	58	22.3	628	12	056253	056253 turnip yell
23	58	22.3	743	5	09N313	09n313 caenorhabdi
24	57.5	22.1	405	16	08BUR5	08uer5 agrobacteri
25	57.5	22.1	500	13	08U009	08uq9 gallus gall
26	57.5	22.1	1451	10	049140	049140 arabidopsis
27	57.5	22.1	1451	10	049142	049142 arabidopsis
28	57.5	22.1	1451	10	049143	049143 arabidopsis
29	57.5	22.1	1675	13	08UUR1	08uuri gallus gall
30	57	21.9	159	13	0919K6	0919k6 xenopus lae
31	56.5	21.7	119	11	09C7L8	09c7l8 mus musculu
32	56.5	21.7	165	11	09CUI9	09cui9 mus musculu
33	56.5	21.7	230	4	08WYK4	08wyk4 homo sapien
34	56.5	21.7	279	4	08WYK3	08wyk3 homo sapien
35	56.5	21.7	431	4	08TDD2	08td22 homo sapien
36	56.5	21.7	480	13	091046	091046 gobius sp.
37	56.5	21.7	714	3	08WZ47	08wz47 neurospora
38	56.5	21.7	807	10	09X107	09x107 arabidopsis
39	56	21.5	228	13	091AV3	091av3 brachydanio
40	55.5	21.3	1329	16	086525	086525 streptomyces
41	55.5	21.3	122	2	P72390	P72390 streptomyces
42	55.5	21.3	267	12	091E26	091e26 potato virus
43	55.5	21.3	359	5	09VJR0	09vjr0 dirosophila
44	55.5	21.3	385	16	09A5B0	09a5b0 caulobacter
45	55.5	21.3	455	16	08Z535	08z535 salmonella

ALIGNMENTS

RESULT 1	096CH9	ID	096CH9	PRELIMINARY;	PRT;	592 AA.
AC	096CH9;	AC	096CH9;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)	DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hypothetical 66.4 kDa protein (Fragment).	DE	Hypothetical 66.4 kDa protein (Fragment).			
OS	Homo sapiens (Human).	OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;	OX	NCBI_TaxID=9606;			
RN	[1]	RN	[1]			
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.			
RC	TISSUE=COLON;	RC	TISSUE=COLON;			
RA	Strausberg R.;	RA	Strausberg R.;			
RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.	RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: BC014227; AAH14227.1; -	DR	EMBL: BC014227; AAH14227.1; -			
FT	KW Hypothetical protein.	FT	KW Hypothetical protein.			
FT	NOR-TER 1	FT	NOR-TER 1			
SQ	SEQUENCE 592 AA; 66400 MW; FA1EB07B0DF47B5 CRC64;	SQ	SEQUENCE 592 AA; 66400 MW; FA1EB07B0DF47B5 CRC64;			
	Query Match		Query Match			
	Best Local Similarity 100.0%; Score 260; DB 4; Length 592;		Best Local Similarity 100.0%; Score 260; DB 4; Length 592;			
	Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0		Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
OY	1 EPLGPRGDSPLQRPQHLMDOGQMRHSFSAGPELLRODKRRPRSGSGTGS 49	OY	1 EPLGPRGDSPLQRPQHLMDOGQMRHSFSAGPELLRODKRRPRSGSGTGS 49			
Db	467 EPLGPRGDSPLQRPQHLMDOGQMRHSFSAGPELLRODKRRPRSGSGTGS 515	Db	467 EPLGPRGDSPLQRPQHLMDOGQMRHSFSAGPELLRODKRRPRSGSGTGS 515			
RESULT 2						
Q9ULH6		ID	Q9ULH6	PRELIMINARY;	PRT;	1807 AA.
AC	Q9ULH6; Q96P46;	AC	Q9ULH6; Q96P46;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	KIA1244 protein (BIG3) (Fragment).	DE	KIA1244 protein (BIG3) (Fragment).			

```

GN K1A1244.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RP TISSUE=BRAIN.
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.;"
RL DNA Res. 6:337-345(1999).
RN (2)
RP SEQUENCE OF 38-1807 FROM N.A.
RN Hong W.,
RA "K1A1244 as a novel distantly related member (B13) of the B1G1/Sec7p
RT subfamily of ANF GEFs.;"
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB033070; BAA86558.1;
DR EMBL; AF413080; AA104174.1;
DR InterPro: IPR000904; Sec7.
DR SMART: SM00222; Sec7; 1.
FT SEQUENCE
FT NCBI_TIER 1
SQ SEQUENCE 1807 AA; 199921 MW; B76C29916F72CB1 CRC64;

Query Match 100.0%; Score 260; DB 4; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2,5e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 EPILPGGDSPLLRPHILMDGGKRNHSFSAGPELLRDKRPRSGSGS 49
Db 1682 EPILPGGDSPLLRPHILMDGGKRNHSFSAGPELLRDKRPRSGSGS 1730

RESULT 3
O9H019 PRELIMINARY; PRT; 462 AA.
ID O9H019;
AC O9H019;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Probable two-component response regulator.
GN PA5166.
GN Pseudomonas aeruginosa.
OS Pseudomonas.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN (1)
RP SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Gabler R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Ladbij K., Lim R.M.,
RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.;"
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.;"
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
CC DOMAIN.
DR EMBL; AE004929; AAC08551.1; -.
DR HSSP; P10958; IDBW.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR002197; HTH_Fis
DR InterPro: IPR001789; Response_reg.
DR InterPro: IPR002078; S1954_Interact.
DR Pfam; PF02954; HTH_8; 1.

```

	Pfam:	PF00712;	response_reg; 1.	
DR	Pfam:	PF00158;	Sigma54_activat; 1.	
DR	PRINTS:	PR01590;	HTHFIS.	
DR	PRODOM:	PD000039;	Response_reg; 1.	
DR	SMART:	SMO0382;	AAA; 1.	
DR	TIGRFAMs:	TIGR01199;	HTH_fis; 1.	
DR	PROSITE:	PS00675;	SIGMA54_INTERACT_1; 1.	
DR	PROSITE:	PS00676;	SIGMA54_INTERACT_2; 1.	
DR	PROSITE:	PS00688;	SIGMA54_INTERACT_3; 1.	
KW	APROSITE:	PS50043;	SIGMA54_INTERACT_4; 1.	
KM	APP-binding:	DNA-binding;	Phosphorylation; Sensory transduction;	
KX	Transcription regulation;	Complete proteome.		
QY	SEQUENCE	462 AA;	50864 MW; 981D8847B2EABBF CRC64;	
Db	Query Match	Best Local Similarity	24.2%; Score 63; DB 16; Length 462;	
	Matches	14; Conservative	3; Mismatches 10; Indels 2; Gaps 1;	
		PLRGGRGDSPLLQRPHILMDQGMRHSFS	30	
		I I :I I I I I I I I I :		
		322 PLRGGEDIPLLFR -HFAEGCARRHOLT	348	
	RESULT 4			
ID	024754	PRELIMINARY;	PRT; 1655 AA.	
AC	024754;			
DT	01-NOV-1996 (TREMblrel. 01,	Created)		
DT	01-NOV-1996 (TREMblrel. 01,	Last sequence update)		
DE	01-JUN-2001 (TREMblrel. 17,	Last annotation update)		
MN	MASTERMIND.			
GN	NAM.			
OS	Drosophila virilis (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OX	Ephydroidea; Drosophilidae; Drosophila.			
RN	[1]			
RA	NEWLINE OF 260-762 FROM N.A.			
RR	Newfield S.J.; Smoller D.A., Yedvobnick B.;	"interspecific comparison of the unusually repetitive drosophilla locus mastermind".		
RL	J. Mol. Evol. 32:415-420(1991).			
RN	[2]			
RA	NEWLINE FROM N.A.			
RR	Newfield S.J.; Schmid A.T., Yedvobnick B.;	"Homopolymer length variation in the Drosophila gene mastermind.";		
RL	J. Mol. Evol. 37:483-495(1993).			
RN	[3]			
RA	NEWLINE FROM N.A.			
RR	Medline=94365848; PubMed=8083889;			
RL	Newfield S.J., Tschida H., Yedvobnick B.;	"Drive-selection equilibrium: homopolymers evolution in the Drosophila gene mastermind."		
RL	J. Mol. Evol. 38:637-641(1994).			
RN	[4]			
RA	NEWLINE FROM N.A.			
RR	Newfield S.J.;			
DL	Submitted (May-1992) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL: M92914; AAC37201.1;	-		
DR	FlyBase: FBgn0013119; Divr\mam.			
SO	SEQUENCE	1655 AA; 175048 MW; 247D8FDFFDDDEL2 CRC64;		
Oy	Query Match	Best Local Similarity	24.0%; Score 62.5; DB 5; Length 1655;	
	Matches	16; Conservative	9; Mismatches 22; Indels 11; Gaps 1;	
		2 PLRGGRGDSPLLQRPHILMDQGMRHSFSAG-----PELLNDKRPARGSGTG	48	
		I I :I I I I I I I I I :		
db	1451 PNGGGOMULTPRXXOMOOLHNHRDOODOOHHNRPGGGGGGGCGGGMOMOODLLDOONAAAGGCg	1508		

RESULT 5

09X6T3 PRELIMINARY; PRT: 293 AA.
 AC 09X6T3; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Glycosyltransferase.
 GN SPFG.
 OS Streptomyces spectabilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=66270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC27741;
 RA Hyun C.G., Suh J.W.;
 RT Isolation of the genes for glycosyltransferase and N-methyltransferase from Streptomyces spectabilis ATCC27741, a spectinomycin producer.
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF145038; AAD28487.1;
 DR InterPro: IPR001173; Glycos_transf_2.
 DR Pfam: PF00535; Glycos_transf_2; 1.
 KW Transferase.
 SQ SEQUENCE 293 AA: 32831 MW: A52A4A0F4B5A5A6B CRC64;

Query Match 23.1%; Score 60; DB 2; Length 293;
 Best Local Similarity 41.7%; Pred. No. 5;
 Matches 15; Conservative 7; Mismatches 12; Indels 2; Gaps 1;
 Oy 1 EPILGPRGDSPLORPOHMDGOMRHSFSGPRL 36
 Db 236 EPMVCTGRDAPL--PGRLLQOPDRHARGAGRDL 269

RESULT 6

09RT11 PRELIMINARY; PRT: 333 AA.
 AC 09RT11; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ADP-ribosylglycohydrolase, putative.
 GN DRI958.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Hatt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vanathavan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI.
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002034; AAF11509.1;
 DR TIGR: DRI958;
 DR InterPro: IPR000209; Peptidase_S8.
 DR PROSITE: PS00136; SUBTILASE ASP; UNKNOWN_1.
 KW Hydrolase; Complete proteome.
 SQ SEQUENCE 333 AA: 34881 MW: 47042A084B5D638 CRC64;

Query Match 23.1%; Score 60; DB 16; Length 333;

Best Local Similarity 34.9%; Pred. No. 5.8;
 Matches 15; Conservative 7; Mismatches 11; Indels 10; Gaps 1;

Oy 4 GPRGDSPLORPOHMDGOMRHSFSG-----PELL 36
 Db 290 GARGLTSPLEVPPEHLLPELRLGHSPGOWORDMACTEHPPELL 332

RESULT 7

09YBM1 PRELIMINARY; PRT: 345 AA.
 AC 09YBM1; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein APE1577.
 GN APE1577.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococcaceae; Aeropyrum.
 NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KI;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamizawa M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix KI.
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000062; BAA80577.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 345 AA: 38129 MW: CF1D95C2B17E92E8 CRC64;

Query Match 22.9%; Score 59.5; DB 17; Length 345;
 Best Local Similarity 31.9%; Pred. No. 7;
 Matches 23; Conservative 8; Mismatches 16; Indels 25; Gaps 5;

Oy 1 EPILGPRGDSPLORPOHMDGOMRHSFSGPRL-----GOM-----RHSGFSGPRL-----LR 37
 Db 244 EPILHRELGDAPGGERPHEVLNPDQLAQSGALGRCLHRRHNVSPPSLSQNALR 303

Oy 38 --QDKPRSGST 47
 Db 304 WPSPTRRKGSS 315

RESULT 8

09RIY2 PRELIMINARY; PRT: 271 AA.
 AC 09RIY2; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hex (Ph) protein.
 GN HHEX OR HEX (PH).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=LIVER;
 RA Myint Z., Inazu T., Tanaka T., Yamada K., Keng V.W., Inoue Y.,
 RA Kuriyama M., Noguchi T.;
 RT Genomic organization and promoter analysis of a mouse homeobox gene, Hex.
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -i- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

[illegible]

RL J. Biol. Chem. 275:20110-20116(2000).
 DR EMBL: AF187891; AAF81404.1; -
 DR InterPro: IPR005135; Exo_endo_phos.
 DR InterPro: IPR000300; IPIC.
 DR InterPro: IPR001230; Prenyl_site.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR SMART: SM00128; IPIC; 1.
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN.1.
 SQ SEQUENCE 644 AA; 70150 MW; 1BB794522C94A49F CRC64;

Query Match 22.7%; Score 59; DB 4; Length 644;
 Best Local Similarity 36.0%; Pred. No. 17;
 Matches 18; Conservative 3; Mismatches 27; Indels 2; Gaps 1;

OY 2 PLGPRGDSPLLRPQHLMDOGMRHSFSAPELLR-QQDKRRSGSTGS 49
 DB 69 PIARPPARPLRLERALSDDKGMRRRRRFRSGOEDLEARNNGTSPSRGSVQS 118

RESULT 12
 O85275 PRELIMINARY: PRT: 276 AA.

ID 085275
 AC 085275.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P1 gene (Fragment).
 GN P1.
 OS Potato virus Y.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus.
 OX NCBI_TaxID=12216;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pehu T., Maki-Valkama T., Valkonen J.P.T., Koivu K., Lehto K.,
 RT "Potato plants transformed with a potato virus Y P1 gene sequence are
 resistant to PVY-O."
 RT Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
 DR EMBL: X82848; CAA58051.1; -
 DR MEROPS: S30.001; -
 DR InterPro: IPR002540; Poty_P1.
 DR Pfam: PF01577; Poty_P1; 1.
 FT NON_TER 276
 SQ SEQUENCE 276 AA; 31269 MW; 00A74C8425D5F2BD CRC64;

Query Match 22.5%; Score 58.5; DB 12; Length 276;
 Best Local Similarity 34.5%; Pred. No. 7.5;
 Matches 19; Conservative 8; Mismatches 15; Indels 13; Gaps 3;

OY 1 EPLGPRG--QDSPLLR---PPQHLMDOGMRHSFSAPELLRQDKRRSGSTGS 49
 DB 117 EPQAPRGIIHTTPVKKYKTRPIIKLTGQMDH-----LIKQVKQIMSGKRGS 164

RESULT 13
 O9C2E4 PRELIMINARY: PRT: 580 AA.

ID 09C2E4
 AC 09C2E4.
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Related to vacuolar protein sorting-associated protein VPS5.
 GN 966.210.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algen V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nykattura G., Meves H.W., Mannhaupt G.,
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.

RN [12]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL513463; CAC28769.2; -
 DR InterPro: IPR001683; PX.
 DR Pfam: PF00787; PX; 1.
 DR SMART: SM00312; PX; 1.
 SQ SEQUENCE 580 AA; 64100 MW; 311F73EF889C7E2F CRC64;

Query Match 22.5%; Score 58.5; DB 3; Length 580;
 Best Local Similarity 32.7%; Pred. No. 17;
 Matches 17; Conservative 7; Mismatches 21; Indels 7; Gaps 2;

OY 1 EPLGPRG-----QDSPLLRPQHLMDOGMRHSFSAPELLRQD-KRRSG 45
 DB 74 DPLGPGASTPAPADTDTPLAPPQPLEQLPLRTLTSGPNOIGSGARPPGGG 125

RESULT 14
 O96MJ2 PRELIMINARY: PRT: 778 AA.

ID 096MJ2
 AC 096MJ2.
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Paromomycin-resistance protein Pm1.
 GN Pm1.
 OS Kluyveromyces fragilis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schmuck E.M., Chen X.J.;
 RT "The yeast Pm1 gene is involved in resistance to paromomycin."
 RT Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF285275; AAK69551.1; -
 SQ SEQUENCE 778 AA; 88129 MW; 6262CEAEAB178BA4 CRC64;

Query Match 22.5%; Score 58.5; DB 3; Length 778;
 Best Local Similarity 32.5%; Pred. No. 24;
 Matches 13; Conservative 6; Mismatches 20; Indels 1; Gaps 1;

OY 2 PLGPRGDSPLLRPQHLMDOGMRHSFSAPELLRQDK 40
 DB 617 PFSPSSQOHHLQQAPOHMASIGSSQSLSPAPQKQSDR 656

RESULT 15

ID 024593 PRELIMINARY: PRT: 3429 AA.

ID 024593
 AC 024593.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE WD-40 repeat protein ANON-X.
 GN CG3585 OR ANON-X OR DMX.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=506RY;
 RX MEDLINE=96398448; PubMed=9729422;
 RA Kraemer C., Well B., Christmann M., Schmidt E.R.;
 RT "The new gene dmX from Drosophila melanogaster encodes a novel WD-
 repeat protein".
 RL Gene 216:267-276(1998).
 CC -1- SIMILARITY: CONTAINS 9 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: X97196; CAA65830.1; -

DR FlyBase: FBgn0023458; CG3585.
DR InterPro: IPR002106; ATRNA_Ligase1.
DR InterPro: IPR000903; NMT.
DR InterPro: IPR00217; Tubulin.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 9.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE: PS00976; NMT_2; UNKNOWN_1.
DR PROSITE: PS00227; TUBULIN; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE: PS50082; WD_REPEATS_2; 1.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 3429 AA; 377671 MW; 4857C939B8A13B9 CMC64;

Query Match	22.5%	Score 58.5	DB 5	Length 3429
Best Local Similarity	41.0%	Pred. No. 1.3e+02		
Matches 16	Conservative 5	Mismatches 11	Indels 7	Gaps 2

```

Oy      9  DSPILQRPQHLMDGQMRHSFSAGPELLRQDKRPPSGST 47
          |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      1334 DSRLLAKQRL--NQVLHTFAHDPE---SRPSRSGET 1365

```

Search completed: June 17, 2003, 11:59:50
Job time : 14.1962 secs

Job time : 14.1962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:16 : Search time 11.5373 Seconds
(without alignments)
461.981 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKYEKKDPSRKEMWENAGNKITYMAADTKISKLMTREK 40

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	215	100.0	40	PS118 prostate mar
2	215	100.0	19	PS118 prostate enco
3	215	100.0	518	PS118 prostate mar
4	215	100.0	1807	Novel human diagno
5	215	100.0	1982	Novel human diagno
6	59.5	27.7	388	Helicobacter pylori
7	57	26.5	664	Mouse oocyte prote
8	56.5	26.3	388	H. pylori GHPD 113
9	56.5	26.3	388	Helicobacter pylori
10	56	26.0	219	Arabidopsis thalia

11	56	26.0	219	21	AA650369	Arabidopsis thalia
12	56	26.0	221	21	AA644031	Arabidopsis thalia
13	56	26.0	221	21	AA650368	Arabidopsis thalia
14	56	26.0	227	21	AA650367	Arabidopsis thalia
15	56	26.0	241	21	AA644030	Arabidopsis thalia
16	56	26.0	304	23	AB891690	Herbicidally activ
17	56	26.0	317	21	AA650357	Arabidopsis thalia
18	56	26.0	326	21	AA650356	Arabidopsis thalia
19	56	26.0	328	21	AA650355	Arabidopsis thalia
20	55	25.6	351	20	AA652796	Arabidopsis thalia
21	55	25.6	654	23	AB893787	Tobacco TORB2 pro
22	53	24.7	141	22	AB604874	Herbicidally activ
23	53	24.7	270	21	AA618382	Novel human diagno
24	53	24.7	338	21	AA618381	Arabidopsis thalia
25	53	24.7	373	23	AAU11287	Arabidopsis thalia
26	53	24.7	410	23	AB893088	Human transducin p
27	53	24.7	423	23	AB893088	Herbicidally activ
28	53	24.7	462	22	AA681158	Myobacterium tube
29	53	24.7	488	22	AB860032	Drosophila melanog
30	53	24.7	528	22	AA693594	Human protein sequ
31	53	24.7	656	23	AA616259	Human kinase PKIN-
32	53	24.7	900	23	AB893716	Herbicidally activ
33	53	24.7	909	22	AAU03554	Human protein kina
34	52.5	24.4	232	21	AA638330	Human secreted pro
35	52.5	24.4	234	22	AA699771	Hevea brasiliensis
36	52.5	24.4	234	22	AA699771	An isopenyl dipho
37	52	24.2	197	20	AA648482	Human breast tumou
38	52	24.2	212	20	AA659805	Human normal ovari
39	52	24.2	216	22	AB860725	Drosophila melanog
40	52	24.2	261	22	AB609146	Novel human diagno
41	52	24.2	270	21	AA649118	Arabidopsis thalia
42	52	24.2	330	21	AA649118	Arabidopsis thalia
43	52	24.2	330	23	AB892353	Herbicidally activ
44	52	24.2	338	21	AA649117	Arabidopsis thalia
45	52	24.2	386	21	AA624042	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AA650812	AA650812 standard; Protein: 40 AA.
XX	XX
AC	AA650812;
XX	XX
DT	01-MAY-2002 (first entry)
XX	XX
DE	PS118 prostate marker immunogenic polypeptide.
XX	XX
KW	PS118; prostate; marker; prostate cancer; tumour; metastasis;
KW	benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW	prostatitis; human; diagnosis; therapy; vaccine; immunogen.
XX	XX
OS	Homo sapiens.
XX	XX
PN	US2001055758-A1.
XX	XX
PD	27-DEC-2001.
XX	XX
FE	23-APR-1998; 98US-0065383.
XX	XX
PR	23-APR-1997; 97US-0842385.
XX	XX
PA	(BULL/) BILLING-MEDEL P A.
PA	(COHE/) COHEN M.
PA	(COP/) COPLITTIS T L.
PA	(FRIE/) FRIEDMAN P N.
PA	(GORD/) GORDON J.
PA	(GRAN/) GRANADOS E N.
PA	(HODG/) HODGES S-C.
PA	(KLAAS/) KLAAS M R.
PA	(KRAT/) KRATOCHVIL J D.

PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Copipits TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvill JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI: 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease.
 XX
 PS Claim 17: Page 44; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 332-371 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC AAB91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis, and
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 CC
 SQ Sequence 40 AA:
 Query Match 100.0%; Score 215; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2,4e-21;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTIISKLTMEYK 40
 ||||||||||||||||||||||||||||||||||||||||
 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTIISKLTMEYK 40
 Db
 RESULT 2
 AAM85472
 ID AAM85472 standard; Protein: 518 AA.
 XX
 AC AAM85472;
 XX
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 XX WO9848054-A1.
 XX
 XX 29-OCT-1998.
 PD
 XX 23-APR-1998; 98WO-US08239.
 PF
 XX 23-APR-1997; 97US-0842385.
 PR
 XX (ABBO) ABBOTT LAB.
 PA
 XX Billing-Medel PA, Cohen M, Copipits TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvill JD;
 PI Robert-Strapp L, Russell JC, Stroupe SD;

XX
 DR WPI: 1998-610000/51.
 DR N-PSDB; AAV82812.
 XX
 PT New P118 nucleic acid and proteins - used for diagnosis and
 PT treatment of prostatic disease, especially cancer, and also for drug
 PT screening
 XX
 PS Claim 17: Page 93-94; 117pp; English.
 XX
 CC The present sequence is encoded by consensus PS118 sequence derived from
 CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
 CC were identified from a cDNA library made from prostate tumour tissue.
 CC Recombinant PS118 protein is used to detect PS118-specific antibodies,
 CC to raise antibodies for detection of PS118 antigens, and to screen for
 CC specific binding agents (potential therapeutics), and to isolate specific
 CC antibodies from serum. Detection of PS118 protein or nucleic acid, which
 CC are prostate related, and altered or elevated in prostatic disease, is
 CC used for detection, diagnosis, staging, monitoring and prognosis of
 CC prostatic disease, particularly cancer, and to identify subjects at
 CC risk.
 CC
 SQ Sequence 518 AA;
 Query Match 100.0%; Score 215; DB 19; Length 518;
 Best Local Similarity 100.0%; Pred. No. 4.3e-20;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTIISKLTMEYK 40
 ||||||||||||||||||||||||||||||||||||||||
 332 SPKVEKKDPSRKKEMENAGNKITYTMAADKTIISKLTMEYK 371
 Db
 RESULT 3
 AAM50809
 ID AAM50809 standard; Protein: 518 AA.
 XX
 AC AAM50809;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker partial sequence.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 XX US2001055758-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 23-APR-1998; 98US-0065383.
 PF
 XX 23-APR-1997; 97US-0842385.
 PR
 XX (BTL/) BILLING-MEDEL P A.
 PA
 XX (COHE/) COHEN M.
 PA (COPL/) COPIPITS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVILL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 XX Billing-Medel PA, Cohen M, Copipits TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvill JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI: 2002-187683/24.

DR N-PSDB: ABA91651.
XX
PT Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease.
XX
PS Claim 17: Page 42-43; 57pp; English.
XX
CC The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus CDNA sequence
CC (see ABA91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.
XX
SQ Sequence 518 AA:
Query Match 100.0%; Score 215; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 4.3e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPVKEKKDPSRKKEWMENAGNKITYTMAADTKISKLMEYK 40
DB 332 SPVKEKKDPSRKKEWMENAGNKITYTMAADTKISKLMEYK 371
RESULT 4
ID ABOG09728 standard; Protein: 1807 AA.
XX
AC ABOG09728;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9719.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS73915.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT Biodiversity
XX
PS Claim 20: SEQ ID NO 40087; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABOG0010-ABG3037 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 1807 AA:
Query Match 100.0%; Score 215; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.7e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPVKEKKDPSRKKEWMENAGNKITYTMAADTKISKLMEYK 40
DB 1621 SPVKEKKDPSRKKEWMENAGNKITYTMAADTKISKLMEYK 1660
RESULT 5
ID ABOG09731 standard; Protein: 1982 AA.
XX
AC ABOG09731;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9722.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS73918.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 4009b; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG3037 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1982 AA;
Query Match 100.0%; Score 215; DB 22; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1,9e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SPKVEKKDPSRKEMWENAGNKITYTMAADTKISLMTYK 40
DB 1735 SPKVEKKDPSRKEMWENAGNKITYTMAADTKISLMTYK 1774
RESULT 6
AAU36030
ID AAU36030 standard; Protein: 388 AA.
XX
AC AAU36030;
XX
DT 14-FEB-2002 (first entry)
XX
DE Helicobacter pylori cellular proliferation protein #343.
XX
KW Antisense: prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Helicobacter pylori.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS3889.
XX
PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11623; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation. Their use in identifying the
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 388 AA;
Query Match 27.7%; Score 59.5; DB 22; Length 388;
Best Local Similarity 37.1%; Pred. No. 13;
Matches 13; Conservative 4; Mismatches 13; Indels 5; Gaps 1;
QY 11 RKKEWMENAGNKITYTMAADK-----TSKLMTEYK 40
DB 217 RLKLMWSNLQNSLFTLLPKLANALRISDLPESIQ 251
RESULT 7
AAE06034
ID AAE06034 standard; Protein: 664 AA.
XX
AC AAE06034;
XX
DT 25-SEP-2001 (first entry)
XX
DE Mouse oocyte protein 5 (MOP5).
XX
KW Mouse oocyte protein 5; MOP5; egg specific surface protein; immunogen;
KW vaccine; fertility; contraceptive; active immunisation; sterilisation.
XX
OS Mus musculus.
XX
PN WO200153339-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US01718.
XX
PR 20-JAN-2000; 2000US-0177123.
XX
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Herr JC, Coonrod SA, Wright P;
XX
DR WPI; 2001-465367/50.
DR N-PSDB; AAD11617, AAD11618.
XX
PT New egg-specific surface proteins, useful as immunogens in a vaccine
PT preparation for modulating fertility, particularly useful for producing
PT antibodies for temporary, reversible contraception methods -
XX
XX Claim 1; Page 32-35; 41pp; English.
XX
CC The present sequence is mouse oocyte protein 5 (MOP5). The
CC present invention relates to egg specific surface proteins

Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 1;
QY 11 RKKEMWENAGNKIYTMADK-----TISKLMTEYK 40
| | | | | : : : | | | :
Db 217 RLKLMMSNLQNSLFTLLPRLANALRISDPESYQ 251
RESULT 10
AAG44032
ID AAG44032 standard; Protein: 219 AA.
XX
AC AAG44032;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55105.
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123538.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR	13-AUG-1999;	99US--0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152365.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

26.0%; Score 56; DB 21; Length 219;

```

Best Local Similarity: 33.38; FREQ: NO: 15;
Matches 10; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

```

```

Oy      9 PSRKKEWENAGNKIYTMADKTIISKLMTE 38
         ||::||  ||  :  ::||:
Db      105 PSLQKQWYSVLGNHDYRGNEAQLSKVLTO 134

```

RESULT 11	
AGS50369	
ID	AGS50369 standard; Protein; 219 AA.
XX	
AC	
AGS50369;	
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 63822.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter
XX	termination sequence.
XX	
OS	Arabidopsis thaliana.
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138009.
PR	10-JUN-1999; 99US-0138504.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	17-JUN-1999; 99US-0139453.
PR	18-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.

```
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145216.
PR 27-JUL-1999; 9905-0145593.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0146389.
PR 04-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.

PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152353.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 05-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 26.0%; Score 56; DB 21; Length 219;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 10; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

OY 9 PSRKEMWENAGNKITYTMAADRTSKLMT 38
    | : | : | | : | : | : | : | : |
Db 105 PSLQKQNTSVLGNHDYRGNWEAQLSKVLTQ 134

RESULT 12
ID AA044031 standard; Protein: 221 AA.
XX AC AA044031;
XX DT 18-OCT-2000 (first entry)
```

XX Arabidopsis thaliana protein fragment seq ID NO: 55104.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.

PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155186.
PR 24-SEP-1999; 990S-0155486.
PR 28-SEP-1999; 990S-0155659.
PR 29-SEP-1999; 990S-0156458.
PR 04-OCT-1999; 990S-0156596.
PR 05-OCT-1999; 990S-0157117.
PR 06-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 26.0%; Score 56; DB 21; Length 221;
Best Local Similarity 33.3%; Pred. No. 20;
Matches 10; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 9 PSRRKKWENAGNKKIYTMADKTISKLMTE 38
||:|:|: || | : ||:|:|:
Db 107 PSLQKQWYSVLGNHDYRCGNVEAQLSKVLTQ 136

RESULT 14
AAG50367
ID AAG50367 standard; Protein: 227 AA.
XX
AC AAG50367;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63820.
XX
DE Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
KW Arabidopsis thaliana.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128224.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158028.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.0%; Score 56; DB 21; Length 227;
Best Local Similarity 33.3%; Pred. No. 20;
Matches 10; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
OY 9 PSRKEWENGNGKIYTPAADKTSKLTME 38
DB 113 PSLQKQWYSLVGNHNDYRGWVEAQLSKVLTQ 142

RESULT 15
AAC44030
ID AAC44030 standard; Protein: 241 AA.
XX
AC AAC44030;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55103.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:56:11 ; Search time 3.6896 Seconds
(without alignments)
318.982 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPYVEKKDPSRKEMENAGNKITYTMADRTISKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	40	US-09-065-383-30	Sequence 30, Appl
2	215	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	35	25.6	351	US-09-431-976-2	Sequence 2, Appl
4	52.5	24.4	234	US-09-296-754-2	Sequence 2, Appl
5	51.5	24.0	489	US-09-986-536-2	Sequence 2, Appl
6	51	23.7	528	US-08-793-229-35	Sequence 35, Appl
7	51	23.7	528	US-09-285-957-35	Sequence 35, Appl
8	51	23.7	4544	US-08-468-486-52	Sequence 32, Appl
9	51	23.7	4544	US-08-469-658-52	Sequence 32, Appl
10	50.5	23.5	489	US-09-545-814-29	Sequence 52, Appl
11	50.5	23.5	995	US-08-673-789-5	Sequence 29, Appl
12	50	23.3	92	US-08-713-939A-79	Sequence 79, Appl
13	50	23.3	92	US-09-036-579-79	Sequence 79, Appl
14	50	23.3	92	US-09-550-374-79	Sequence 79, Appl
15	50	23.3	291	US-08-568-459A-13	Sequence 13, Appl
16	50	23.3	291	US-08-487-826B-25	Sequence 25, Appl
17	50	23.3	291	US-09-210-288-13	Sequence 2, Appl
18	50	23.3	397	US-08-774-104A-2	Sequence 2, Appl
19	50	23.3	1115	US-08-568-459A-2	Sequence 2, Appl
20	50	23.3	1115	US-08-487-826B-2	Sequence 2, Appl
21	50	23.3	1115	US-09-210-288-2	Sequence 2, Appl
22	50	23.3	1115	US-09-431-976-2	Sequence 2, Appl
23	49	22.8	476	US-09-346-408-12	Sequence 12, Appl
24	49	22.8	970	US-08-673-789-7	Sequence 7, Appl
25	49	22.8	973	US-08-162-809-8	Sequence 8, Appl
26	49	22.8	973	US-08-162-809-10	Sequence 10, Appl
27	49	22.8	988	US-08-162-809-14	Sequence 14, Appl

28	49	22.8	995	1	US-08-162-809-18
29	49	22.8	1011	1	US-08-162-809-12
30	48.5	22.6	261	1	US-08-624-125-18
31	48.5	22.6	569	1	US-08-306-231-3
32	48.5	22.6	672	3	US-09-040-843-4
33	48.5	22.6	672	3	US-09-040-843-4
34	48.5	22.6	866	3	US-09-621-855-4
35	48.5	22.6	866	3	US-09-621-855-2
36	48.5	22.6	896	2	US-08-640-389A-10
37	48.5	22.6	896	2	US-08-618-957A-10
38	48.5	22.6	898	2	US-08-693-697-36
39	48.5	22.6	898	4	US-08-588-189-3
40	48.5	22.6	906	2	US-08-640-389A-9
41	48.5	22.6	906	4	US-08-618-957A-9
42	48.5	22.6	908	2	US-08-693-697-33
43	48.5	22.6	908	2	US-08-588-526-3
44	48.5	22.6	958	2	US-08-640-389A-8
45	48.5	22.6	958	4	US-08-618-957A-8

ALIGNMENTS

RESULT 1
US-09-065-383-30
Sequence 30, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065.383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842.385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084. US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-30

Query Match 100.0%; Score 215; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 4.3e-23;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTSKLTMEYK 40
Db 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTSKLTMEYK 40

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 215; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 9.7e-22;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPKVEKKDPSRKKEMENAGNKITYTMAADKTSKLTMEYK 40
Db 332 SPKVEKKDPSRKKEMENAGNKITYTMAADKTSKLTMEYK 371

RESULT 3
US-09-431-976-2
Sequence 2, Application US/09431976
Patent No. 6423520

GENERAL INFORMATION:
APPLICANT: Conkling, Mark
APPLICANT: Mendu, Nandini
APPLICANT: Song, Wen
TITLE OF INVENTION: Regulation of Quinolinate Phosphoribosyl Transferase
TITLE OF INVENTION: Expression
FILE REFERENCE: 5051-338
CURRENT APPLICATION NUMBER: US/09/431,976
CURRENT FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/021,286
PRIOR FILING DATE: 1998-02-10
PRIOR APPLICATION NUMBER: 60/049,471
PRIOR FILING DATE: 1997-06-12
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 351
TYPE: PRT
ORGANISM: Nicotiana tabacum
US-09-431-976-2

Query Match 25.6%; Score 55; DB 4; Length 351;
Best Local Similarity 52.9%; Pred. No. 8;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 8 DPSRKKEMENAGNKIY 24
Db 106 DPSLKEVYVNDGDKVH 122

RESULT 4
US-09-296-754-2
Sequence 2, Application US/09296754A
Patent No. 6316695

GENERAL INFORMATION:
APPLICANT: Korea Kumho Petrochemical Co. Ltd.
TITLE OF INVENTION: Isopentenyl diphosphate isomerase from Hevea Brasiliensis
FILE REFERENCE: PX99102/US
CURRENT APPLICATION NUMBER: US/09/296,754A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 2
SOFTWARE: KOPATIN 1.0
SEQ ID NO 2
LENGTH: 234
TYPE: PRT
ORGANISM: IPIHb (IPP isomerase from Hevea brasiliensis)
US-09-296-754-2

Query Match 24.4%; Score 52.5; DB 4; Length 234;
Best Local Similarity 52.2%; Pred. No. 11;
Matches 12; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

OY 14 EWMENAGNKITYTMAAD-KTISKL 35
Db 211 KWMENVENGLKEAVDMKTIKHL 233

RESULT 5
US-09-986-536-2
Sequence 2, Application US/09986536
Patent No. 6461841
GENERAL INFORMATION:

APPLICANT: GEUEKE, BIRGIT
APPLICANT: HUMMEL, WERNER
APPLICANT: BOMMARTUS, ANDREAS
TITLE OF INVENTION: L-Amino Acid Oxidase from Rhodococcus Species
FILE REFERENCE: 215209050X
CURRENT APPLICATION NUMBER: US/09/986, 536
CURRENT FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: DE 100 55 512.8
PRIOR FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 489
TYPE: PRT
ORGANISM: Rhodococcus opacus
US-09-986-536-2

Query Match 24.0%; Score 51.5; DB 4; Length 489;
Best Local Similarity 35.6%; Pred. No. 37;
Matches 16; Conservative 6; Mismatches 14; Indels 9; Gaps 3;

OY 3 KVEKKDPSRK-----KEMENAGNKITYTMAA--DKTISKLMTEY 39
DB 313 KAKPSSSGKLGIEYSRRMETE-DRYGGASNTDKDISQIMF 356

RESULT 6
US-08-793-229-35
Sequence 35, Application US/08793229
Patent No. 5891703

GENERAL INFORMATION:
APPLICANT: VAN DER LAAN, Jan Metske
APPLICANT: RIEMENS, Adriana Marina
APPLICANT: QUAX, Wilhelmus Johannes
TITLE OF INVENTION: Mutated Penicillin G Acylase Genes
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,229
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/03249

ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)913-0001
TELEFAX: (312)913-0002
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-793-229-35

Query Match 23.7%; Score 51; DB 2; Length 528;
Best Local Similarity 27.7%; Pred. No. 47;
Matches 13; Conservative 3; Mismatches 11; Indels 20; Gaps 1;

OY 14 EWMENAGNKIY-----TMAADKTIISKLMTEYK 40
DB 355 EWMNNLHDKLFEMDELGFYGTKEITDHRYGASLAYKNISKESPTNYK 401

RESULT 7
US-09-285-957-35
Sequence 35, Application US/09285957
Patent No. 603823

GENERAL INFORMATION:
APPLICANT: VAN DER LAAN, Jan Metske
APPLICANT: RIEMENS, Adriana Marina
APPLICANT: QUAX, Wilhelmus Johannes
TITLE OF INVENTION: Mutated Penicillin G Acylase Genes
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/285,957
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/793,229

ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)913-0001
TELEFAX: (312)913-0002
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-285-957-35

Query Match 23.7%; Score 51; DB 3; Length 528;
Best Local Similarity 27.7%; Pred. No. 47;
Matches 13; Conservative 3; Mismatches 11; Indels 20; Gaps 1;

OY 14 EWMENAGNKIY-----TMAADKTIISKLMTEYK 40
DB 355 EWMNNLHDKLFEMDELGFYGTKEITDHRYGASLAYKNISKESPTNYK 401

RESULT 8
US-08-469-486-52
Sequence 52, Application US/08469486
Patent No. 5739281
GENERAL INFORMATION:
APPLICANT: Thøgersen, Hans Christian
APPLICANT: Hollet, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: Improved method for the refolding of
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts

Query Match 23.7%; Score 51; DB 3; Length 528;
Best Local Similarity 27.7%; Pred. No. 47;
Matches 13; Conservative 3; Mismatches 11; Indels 20; Gaps 1;

```

ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-658-52

Query Match
Best local Similarity 23.7%; Score 51; DB 2; Length 4544;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1

OY 8 DPSRKKEWENNAGN--KIYMAADKTIKLTME 38
| | | | | | | | | | | | | | | | | |
Db 4009 DPLRGTWSDGWNHPKIEIETAAIDGTLRETLV 4041

RESULT 10
US-09-545-814-29
Sequence 29, Application US/09545814
Patent No. 6416977
GENERAL INFORMATION:
APPLICANT: Beecher, Anna M.
TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
FILE REFERENCE: FC-5-C1
CURRENT APPLICATION NUMBER: US/09/545,814
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 60/128,833
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 489
TYPE: PRT
ORGANISM: Tagged Ctenocephalides felis
US-09-545-814-29

Query Match
Best local Similarity 23.5%; Score 50.5; DB 4; Length 489;
Matches 14; Conservative 6; Mismatches 16; Indels 11; Gaps 2

OY 5 EKKDSRKK---EWMENAGNKIYMAADK-----TIKLTMEYK 40
| | | | | | | | | | | | | | | | | |
Db 29 DDKDPVKLIIVAGWAEKGKRYSTMAVEKKRRAFIKRVYDFNNYK 75

RESULT 11
US-08-673-789-5
Sequence 5, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, REMPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK

```

STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,789
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 995
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-5

Query Match 23.5%; Score 50.5; DB 2; Length 995;
Best Local Similarity 37.1%; Pred. No. 1.2e+02;
Matches 13; Conservative 5; Mismatches 8; Indels 9; Gaps 1;

OY 15 WENAGNKIYTMADKTI-----KLMTEYK 40
Db 138 WMENPMKVDITIADESFQVDLGRVVKLMTVEYR 172

RESULT 12
US-08-713-939A-79
Sequence 79, Application US/08713939A
Patent No. 5846533
GENERAL INFORMATION:
APPLICANT: Prusiner, Stanley B.
APPLICANT: Williamson, R. Anthony
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PRP
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,939A
FILING DATE: 13-SEP-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bozicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 06510/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-5277
TELEFAX: 415-854-0875
TELEX:
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-036-579-79

Query Match 23.3%; Score 50; DB 4; Length 92;
Best Local Similarity 29.8%; Pred. No. 7.8;
Matches 14; Conservative 6; Mismatches 15; Indels 12; Gaps 2;

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-5277
TELEFAX: 415-854-0875
TELEX:
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-713-939A-79

Query Match 23.3%; Score 50; DB 2; Length 92;
Best Local Similarity 29.8%; Pred. No. 7.8;
Matches 14; Conservative 6; Mismatches 15; Indels 12; Gaps 2;

OY 6 KKDSRRKEW-----WENAGNKIY-----TMADKTIKLMTEYK 40
Db 13 KORTGQDLEWIGETIPRSGNTYNEKFKGKATLADKSSSTAYMELR 59

RESULT 13
US-09-036-579-79
Sequence 79, Application US/09036579
Patent No. 6290554
GENERAL INFORMATION:
APPLICANT: Prusiner, Stanley B.
APPLICANT: Williamson, R. Anthony
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PRP
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,579
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/713,939
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bozicevic, Karl
REGISTRATION NUMBER: 28,807
REFERENCE/DOCKET NUMBER: 06510/059001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-5277
TELEFAX: 415-854-0875
TELEX:
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-036-579-79

Query Match 23.3%; Score 50; DB 4; Length 92;
Best Local Similarity 29.8%; Pred. No. 7.8;
Matches 14; Conservative 6; Mismatches 15; Indels 12; Gaps 2;

OY 6 KKDSRRKEW-----WENAGNKIY-----TMADKTIKLMTEYK 40

Db 13 KÖRTGÖLEWIGETIPRSGNTYYNEKEKGKATLADSSSTAYMELR 59

RESULT 14

US-09-550-374-79
Sequence 79, Application US/09550374
Patent No. 6372214

GENERAL INFORMATION:

APPLICANT: Prusiner, Stanley B.
APPLICANT: Williamson, R. Anthony
APPLICANT: Burton, Dennis R.
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: U.S.A.

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/550,374

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/036,579

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Bozicevic, Karl

REGISTRATION NUMBER: 28,807

REFERENCE/DOCKET NUMBER: 06510/059001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-5277

TELEFAX: 415-854-0875

TELEX:

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 92 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-550-374-79

Query Match 23.3%, Score 50; DB 4; Length 92;

Best Local Similarity 29.8%; Pred. No. 7.8;

Matches 14; Conservative 6; Mismatches 15; Indels 12; Gaps 2;

Db 13 KÖRTGÖLEWIGETIPRSGNTYYNEKEKGKATLADSSSTAYMELR 59

RESULT 15

US-08-568-459A-13
Sequence 13, Application US/08568459A
Patent No. 5649306

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellem, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbie Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/568,459A

FILING DATE: 07-DEC-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 291 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

US-08-568-459A-13

Query Match 23.3%, Score 50; DB 2; Length 291;

Best Local Similarity 33.3%; Pred. No. 31;

Matches 7; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 123 DEKAQQRKRWNNESKAOIWT 143

Search completed: June 17, 2003, 12:03:40

Job time : 4.6896 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:36 ; Search time 4.27526 Seconds
(without alignments)
1000.321 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPVKEKKDPSRKKEWMENAGNKITYTMADKISKLMTEYK 40

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Published_Applications_AA:*
2: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	27.7	388	10 US-09-815-242-11623	Sequence 11623, A
2	56.5	26.3	388	10 US-09-815-242-11458	Sequence 11458, A
3	56.5	26.3	388	10 US-09-881-752A-350	Sequence 350, App
4	55	25.6	351	10 US-09-963-340-2	Sequence 2, Appl
5	53	24.7	462	9 US-09-712-363-209	Sequence 209, App
6	52.5	24.4	231	9 US-10-050-704-106	Sequence 106, App
7	52	24.2	336	9 US-10-160-293-4	Sequence 4, Appl
8	52	24.2	396	9 US-10-160-293-2	Sequence 2, Appl
9	52	24.2	468	9 US-10-077-040-1	Sequence 1, Appl
10	52	24.2	468	9 US-10-063-541-150	Sequence 150, App
11	52	24.2	468	9 US-10-036-041-67	Sequence 67, App
12	52	24.2	468	9 US-10-028-072-90	Sequence 90, Appl
13	52	24.2	468	9 US-10-035-855-67	Sequence 67, Appl
14	52	24.2	468	9 US-10-063-616-150	Sequence 150, App
15	52	24.2	468	9 US-10-063-502-150	Sequence 90, App
16	52	24.2	468	9 US-10-121-049-90	Sequence 90, Appl
17	52	24.2	468	9 US-10-123-904-90	Sequence 90, Appl
18	52	24.2	468	9 US-10-140-470-90	Sequence 90, Appl
19	52	24.2	468	9 US-09-931-836-67	Sequence 67, Appl

20	52	24.2	468	9	US-10-175-746-90	Sequence 90, Appl
21	52	24.2	468	9	US-10-176-918-90	Sequence 90, Appl
22	52	24.2	468	9	US-10-176-921-90	Sequence 90, Appl
23	52	24.2	468	9	US-10-036-214-67	Sequence 67, Appl
24	52	24.2	468	9	US-10-137-865-90	Sequence 90, Appl
25	52	24.2	468	9	US-10-140-474-90	Sequence 90, Appl
26	52	24.2	468	9	US-10-035-719-67	Sequence 67, Appl
27	52	24.2	468	9	US-10-142-431-90	Sequence 90, Appl
28	52	24.2	468	9	US-10-143-114-90	Sequence 90, Appl
29	52	24.2	468	9	US-10-140-002-90	Sequence 90, Appl
30	52	24.2	468	9	US-10-036-160-67	Sequence 67, Appl
31	52	24.2	468	9	US-10-142-419-90	Sequence 90, Appl
32	52	24.2	468	9	US-10-035-958-67	Sequence 67, Appl
33	52	24.2	468	9	US-10-036-150-67	Sequence 67, Appl
34	52	24.2	468	9	US-10-063-518-150	Sequence 150, App
35	52	24.2	468	9	US-10-063-598-150	Sequence 150, App
36	52	24.2	468	9	US-10-123-262-90	Sequence 90, Appl
37	52	24.2	468	9	US-10-142-423-90	Sequence 90, Appl
38	52	24.2	468	9	US-10-227-693-150	Sequence 150, App
39	52	24.2	468	9	US-10-121-050-90	Sequence 90, Appl
40	52	24.2	468	9	US-10-141-755-90	Sequence 90, Appl
41	52	24.2	468	9	US-10-143-032-90	Sequence 90, Appl
42	52	24.2	468	9	US-10-063-367-150	Sequence 150, App
43	52	24.2	468	9	US-10-123-108-90	Sequence 90, Appl
44	52	24.2	468	9	US-10-123-236-90	Sequence 90, Appl
45	52	24.2	468	9	US-10-123-261-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1

Sequence 11623, Application US/09815242
Patent No. US2002061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: EITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11623

LENGTH: 388

TYPE: PRT

ORGANISM: Helicobacter pylori

US-09-815-242-11623

Query Match

Best Local Similarity

27.7%; Score 59.5; DB 10; Length 388;

37.1%; Pred. No. 9.1;

Matches 13: Conservative 4: Mismatches 13: Indels 5: Gaps 1;
 Qy 11 RKKEWENAGNKITYTMAADK-----TISKLTETK 40
 Db 217 RLKLMWSNLQNSLTLLPDLRLANLRLSDLPESYQ 251

RESULT 2

US-09-815-242-11458
 ; Sequence 11458, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:

APPLICANT: Haseibeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.

TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 11458
 LENGTH: 388
 TYPE: PRT
 ORGANISM: Helicobacter pylori

FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(388)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-815-242-11458

Query Match 26.3%, Score 56.5, DB 10; Length 388;
 Best Local Similarity 34.3%, Pred. No. 22;
 Matches 12: Conservative 5; Mismatches 13; Indels 5; Gaps 1;

Qy 11 RKKEWENAGNKITYTMAADK-----TISKLTETK 40
 Db 217 RLKLMWSNLQNSLTLLPDLRLANLRLSDLPESYQ 251

RESULT 3

US-09-881-752A-350
 ; Sequence 350, Application US/09881752A
 ; Patent No. US20020115078A1
 ; GENERAL INFORMATION:

APPLICANT: Kleantous, Harold
 APPLICANT: Al-Garawi, Amal
 APPLICANT: Miller, Charles
 APPLICANT: Tomb, Jean-Francois

TITLE OF INVENTION: Identification of Polynucleotides
 TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
 TITLE OF INVENTION: Genome

FILE REFERENCE: 06132/041002
 CURRENT APPLICATION NUMBER: US/09/881,752A
 CURRENT FILING DATE: 2001-06-15
 PRIOR APPLICATION NUMBER: US 08/833,457
 PRIOR FILING DATE: 1997-04-01
 NUMBER OF SEQ ID NOS: 370
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 350
 LENGTH: 388
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-09-881-752A-350

Query Match 26.3%, Score 56.5, DB 10; Length 388;
 Best Local Similarity 34.3%, Pred. No. 22;
 Matches 12: Conservative 5; Mismatches 13; Indels 5; Gaps 1;

Qy 11 RKKEWENAGNKITYTMAADK-----TISKLTETK 40
 Db 217 RLKLMWSNLQNSLTLLPDLRLANLRLSDLPESYQ 251

RESULT 4:

US-09-963-340-2
 ; Sequence 2, Application US/09963340
 ; Patent No. US20020108151A1
 ; GENERAL INFORMATION:

APPLICANT: Conkling, Mark
 APPLICANT: Mendu, Nandini
 APPLICANT: Song, Wen
 TITLE OF INVENTION: Regulation of Guinolate Phosphoribosyl Transferase
 FILE REFERENCE: 5051-338
 CURRENT APPLICATION NUMBER: US/09/963,340
 CURRENT FILING DATE: 2001-09-24
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/021,286
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-10
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 351
 TYPE: PRT
 ORGANISM: Nicotiana tabacum
 US-09-963-340-2

Query Match 25.6%, Score 55; DB 10; Length 351;
 Best Local Similarity 52.9%, Pred. No. 31;
 Matches 9: Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 8 DPSRKKEWENAGNKITY 24
 Db 106 DPSLKEVENVNDGDKVH 122

RESULT 5

US-09-712-363-209
 ; Sequence 209, Application US/09712363
 ; Patent No. US20020164588A1
 ; GENERAL INFORMATION:

APPLICANT: Eisenberg, David
 APPLICANT: Rotstein, Sergio H.
 APPLICANT: Marcotte, Edward M.

TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
 FILE REFERENCE: 07419-032001
 CURRENT APPLICATION NUMBER: US/09/712,363
 PRIOR FILING DATE: 2000-11-13
 PRIOR APPLICATION NUMBER: PCT/US00/02246
 PRIOR FILING DATE: 2000-01-28
 PRIOR APPLICATION NUMBER: 60/179,531
 PRIOR FILING DATE: 2000-02-01
 PRIOR APPLICATION NUMBER: 60/117,844
 PRIOR FILING DATE: 1999-01-25

;; PRIOR APPLICATION NUMBER: 60/118,206,
;; PRIOR FILING DATE: 1999-02-01
;; PRIOR APPLICATION NUMBER: 60/126,593
;; PRIOR FILING DATE: 1999-03-26
;; PRIOR APPLICATION NUMBER: 60/134,093
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/134,092
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/165,124
;; PRIOR FILING DATE: 1999-11-12
;; PRIOR APPLICATION NUMBER: 60/165,086
;; PRIOR FILING DATE: 1999-11-12
;; NUMBER OF SEQ ID NOS: 292
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 209
;; LENGTH: 462
;; TYPE: PRT
;; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-209

Query Match 24.7%; Score 53; DB 9; Length 462;
Best Local Similarity 31.6%; Pred. NO. 75;
Matches 12; Conservative 8; Mismatches 16; Indels 2; Gaps 1;

Oy 1 SPKYEKKDPSRKKEWENAGNKIYTMADKTSKLTMT 38
Db 179 APTVAGDVR--AMWDLAGNRAPPSIARAVSKVIAE 214

RESULT 6
US-10-050-704-106
;; Sequence 106, Application US/10050704
;; Publication No. US20030050442a1
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: 62 Human Secreted Proteins
;; FILE REFERENCE: P2039P1
;; CURRENT APPLICATION NUMBER: US/10/050,704
;; CURRENT FILING DATE: 2002-01-18
;; PRIOR APPLICATION NUMBER: 09/684,524
;; PRIOR FILING DATE: 2000-10-10
;; PRIOR APPLICATION NUMBER: PCT/US00/08979
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: 60/128,693
;; PRIOR FILING DATE: 1999-04-09
;; PRIOR APPLICATION NUMBER: 60/130,991
;; PRIOR FILING DATE: 1999-04-26
;; NUMBER OF SEQ ID NOS: 344
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 106
;; LENGTH: 231
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-050-704-106

US-10-050-704-106

Query Match 24.4%; Score 52.5; DB 9; Length 231;
Best Local Similarity 30.0%; Pred. NO. 40;
Matches 12; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

Oy 4 VEKDPKSRK-----KEMWENAGNKIYTMADK 30
Db 182 IRKKQSSPLANTYINIFIMDKTWMHNSSEKNFTKLAKK 221

RESULT 7
US-10-160-293-4
;; Sequence 4, Application US/10160293
;; Publication No. US20030022208a1
;; GENERAL INFORMATION:
;; APPLICANT: LINDQUIST, Erika et al.
;; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;; TITLE OF INVENTION: USES THEREOF

;; FILE REFERENCE: CLO01241-PROV
;; CURRENT APPLICATION NUMBER: US/10/160,293
;; CURRENT FILING DATE: 2002-06-04
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 336
;; TYPE: PRT
;; ORGANISM: Human
US-10-160-293-4

Query Match 24.2%; Score 52; DB 9; Length 336;
Best Local Similarity 32.6%; Pred. NO. 70;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

Oy 6 KKDPKSRKKEWEN-AG--NKIYTMADK-----DKTISKLTMT 37
Db 77 EEDPNKPTSWTENQAGKIPEKVTPTMAIODGLANGENDETVSNLT 122

RESULT 8
US-10-160-293-2
;; Sequence 2, Application US/10160293
;; Publication No. US20030022208a1
;; GENERAL INFORMATION:
;; APPLICANT: LINDQUIST, Erika et al.
;; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;; TITLE OF INVENTION: USES THEREOF
;; FILE REFERENCE: CLO01241-PROV
;; CURRENT APPLICATION NUMBER: US/10/160,293
;; CURRENT FILING DATE: 2002-06-04
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 396
;; TYPE: PRT
;; ORGANISM: Human
US-10-160-293-2

US-10-160-293-2

Query Match 24.2%; Score 52; DB 9; Length 396;
Best Local Similarity 32.6%; Pred. NO. 84;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

Oy 6 KKDPKSRKKEWEN-AG--NKIYTMADK-----DKTISKLTMT 37
Db 137 EEDPNKPTSWTENQAGKIPEKVTPTMAIODGLANGENDETVSNLT 182

RESULT 9
US-10-077-040-1
;; Sequence 1, Application US/10077040
;; Patent No. US20020156014a1
;; GENERAL INFORMATION:
;; APPLICANT: Lal, Preeti
;; Corley, Neil C.
;; Patterson, Chandra
;; TITLE OF INVENTION: HUMAN NEUROSECRETORY PROTEINS
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Dr.
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/077,040

US-10-077-040-1

```

1      FILING DATE: 14-Feb-2002
2      CLASSIFICATION: <Unknown>
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: 09/062,601
5      FILING DATE: <Unknown>
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Cerrone, Michael C.
8      REGISTRATION NUMBER: 39,132
9      REFERENCE/DOCKET NUMBER: PF-0510 US
10     TELECOMMUNICATION INFORMATION:
11     TELEPHONE: 650-855-0555
12     TELEFAX: 650-845-4166
13     TELEX: <Unknown>
14     INFORMATION FOR SEQ ID NO: 1:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH: 468 amino acids
17     TYPE: amino acid
18     STRADEDNESS: single
19     TOPOLOGY: linear
20     IMMEDIATE SOURCE:
21     LIBRARY: ISLTN0701
22     CLONE: 2379427
23     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
24     -US-10-077-040-1

```

```

: PRIOR APPLICATION NUMBER: 09/380142
: PRIOR FILING DATE: 1999-08-25
: PRIOR APPLICATION NUMBER: 09/644848
: PRIOR FILING DATE: 2000-08-22
: PRIOR APPLICATION NUMBER: 09/747259
: PRIOR FILING DATE: 2000-12-20
: PRIOR APPLICATION NUMBER: 09/816744
: PRIOR FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: 09/854208
: PRIOR FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: 09/854280
: PRIOR FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: 09/874503
: PRIOR FILING DATE: 2001-06-05
: PRIOR APPLICATION NUMBER: 09/869599
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: 09/908,827
: PRIOR FILING DATE: 2001-07-18
: PRIOR APPLICATION NUMBER: PCT/US99/10733
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: PCT/US99/28551
: PRIOR FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/30720
: PRIOR FILING DATE: 1999-12-22
: PRIOR APPLICATION NUMBER: PCT/US00/05601
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: 2000-03-02
: PRIOR APPLICATION NUMBER: PCT/US00/14042
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: PCT/US00/15264
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US00/23522
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: PCT/US00/23328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: 2000-12-01
: PRIOR APPLICATION NUMBER: PCT/US00/34956
: PRIOR FILING DATE: 2000-12-20
: PRIOR APPLICATION NUMBER: PCT/US01/06520
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: PCT/US01/17800
: PRIOR FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: PCT/US01/19692
: PRIOR FILING DATE: 2001-06-20
: PRIOR APPLICATION NUMBER: PCT/US01/21066
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: PCT/US01/21735
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 80
: SEQ ID NO 67
: LENGTH: 468
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-036-041-67

Query Match      24.2%  Score 52:  DB 9:  Length 468:
Best Local Similarity 32.6%  Pred. No. 1e+02:
Matches 15:  Conservative 8:  Mismatches 9:  Indels 14:  Gaps 3:

Qy      6  KNDPSRKKEWMEN-AG---NKITYTMA-----DKTISKLMY 37
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      209 EEDPKKPTSWENQAGKIPEKVTYPMMAIQDGLAKGENDETYSNTL 254

RESULT 12
US-10-028-072-90
: Sequence 90, Application US/10028072
: Publication No. US20030004311A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
```

```

: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tunas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang
: TITLE OF INVENTION:
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/10/028,072
: PRIOR FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: 60/049911
: PRIOR FILING DATE: 1997-06-18
: PRIOR APPLICATION NUMBER: 60/056974
: PRIOR FILING DATE: 1997-08-26
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059115
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059117
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059122
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059184
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/059263
: PRIOR FILING DATE: 1997-09-18
: PRIOR APPLICATION NUMBER: 60/059352
: PRIOR FILING DATE: 1997-09-19
: PRIOR APPLICATION NUMBER: 60/059588
: PRIOR FILING DATE: 1997-09-19
: PRIOR APPLICATION NUMBER: 60/059836
: PRIOR FILING DATE: 1997-09-24
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062285
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062287
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/062814
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/062816
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063045
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063082
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/063127
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/063327
: PRIOR FILING DATE: 1997-10-27
: PRIOR APPLICATION NUMBER: 60/063329
: PRIOR FILING DATE: 1997-10-27
: PRIOR APPLICATION NUMBER: 60/063550
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063661
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/063704
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063733
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063735
: PRIOR FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: 60/063738
: PRIOR FILING DATE: 1997-10-29
```

PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338

PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 24.2% Score 52; DB 9; Length 468;
Best Local Similarity 32.6% Pred. No. 1e+02;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

QY 6 KKDSRKKEWMEN-AG---NKIYMAA-----DKTISKMT 37
DB 209 EDPWKPTSWTENOAGKIPEKVTWPAALQDGLANGENDVTSNLT 254

US-10-035-855-67
Sequence 67, Application US/10035855
Publication No. US20030008348A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3030R1C4
CURRENT FILING DATE: 2001-12-26
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280

PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 67
LENGTH: 468
TYPE: PRT
ORGANISM: Homo Sapien
US-10-035-855-67

Query Match 24.2%; Score 52; DB 9; Length 468;
Best Local Similarity 32.6%; Pred. No. 1e+02;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

QY 6 KKDSRRKKEMWEN-AG---NKIYPMIA-----DKTISKLMT 37
DB 209 EDPKPTSWTENGKIPKVTYPMIAIQDGLANGENDETVSNITLF 254

RESULT 14
US-10-063-616-150
Sequence 150. Application US/10063616
Publication No. US20030013855A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,616

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:06 ; Search time 4.62665 Seconds
(without alignments)
831.137 Million cell updates/sec

Title: US-09-991-681-30
Perfect score: 215
Sequence: 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	29.8	621	1	endo-1,4-beta-xyla
2	63.5	29.5	604	2	DEAD box ATP-depen
3	59.5	27.7	388	2	probable rod shape
4	59	27.4	277	2	oxidoreductase yaf
5	58.5	27.2	104	2	signal recognition
6	57.5	26.7	921	2	isoleucyl-tRNA syn
7	56.5	26.3	158	2	hypothetical prote
8	56.5	26.3	201	2	hypothetical prote
9	56.5	26.3	388	2	cell division prot
10	56.5	26.0	304	1	probable acid phos
11	56	26.0	450	1	signal recognition
12	55	25.6	177	2	protein BYJ6 - com
13	55	25.6	177	2	probable membrane
14	54	25.1	387	2	Ca2+-transporting
15	54	25.1	1048	2	phosphoenolpyruvat
16	53.5	24.9	490	2	phosphoenolpyruvat
17	53.5	24.9	884	2	hypothetical prote
18	53	24.7	259	2	hypothetical prote
19	53	24.7	410	2	restriction modifi
20	53	24.7	425	2	hypothetical prote
21	53	24.7	446	2	signal recognition
22	53	24.7	462	2	probable lipid prot
23	53	24.7	729	2	hypothetical prote
24	52.5	24.4	1289	2	conserved hypotet
25	52	24.2	277	2	hypothetical prote
26	52	24.2	433	2	hypothetical prote
27	52	24.2	599	2	malate-like prote
28	52	24.2	840	2	leucine-tRNA ligas
29	52	24.2	1876	2	zinc metalloprotei

30	52	24.2	1984	2	A44396	P-type cation tran
31	52	24.2	26926	1	I13834	titin, cardiac mus
32	51.5	24.0	318	2	T13015	phosphoprotein pho
33	51.5	24.0	338	2	A82890	hypothetical prote
34	51.5	24.0	921	2	AC1327	isoleucyl-tRNA syn
35	51	23.7	149	2	B96965	hypothetical prote
36	51	23.7	205	2	AE2442	hypothetical prote
37	51	23.7	230	2	CE4396	precortin-2 methyl
38	51	23.7	238	2	A86160	hypothetical prote
39	51	23.7	342	2	E70109	hypothetical prote
40	51	23.7	547	2	A56575	putf-specific nucl
41	51	23.7	572	2	S73730	MG307 homolog H08
42	51	23.7	727	2	T24284	hypothetical prote
43	51	23.7	802	2	I39665	penicillin amidase
44	51	23.7	1017	2	T48452	hypothetical prote
45	51	23.7	1345	2	H90975	hypothetical prote

ALIGNMENTS

RESULT 1
S59632
endo-1,4-beta-xylanase (EC 3.2.1.8) B precursor - Cellvibrio mixtus
C:Species: Cellvibrio mixtus
C:Date: 15-Feb-1996 #sequence_revision 11-Apr-1997 #text_change 18-Jun-1999
C:Accession: S59632; S52742
R:Millward-Sadler, S.J.; Davidson, K.; Hazlewood, G.P.; Black, G.W.; Gilbert, H.J.; C
Biochem. J. 312, 39-48, 1995
A:Title: Novel cellulose-binding domains, NodB homologues and conserved modular archi
A:Reference number: S59631; MUID:96077124; PMID:7492333
A:Accession: S59632
A:Molecule type: DNA
A:Residues: 1-621 <MIL>
A:Cross-references: EMBL:Z48926; NID:q757808; PIDN:CNA88762.1; PID:q757809
C:Genetics:
A:Gene: xynB
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Pseudomonas endo-1,4-beta-xylanase F; Streptomyces endo-1,4-beta-xylan
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-621/Product: endo-1,4-beta-xylanase B #status predicted <MAT>
F:302-615/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXT>
F:403,516/Active site: Glu #status predicted

Query Match 29.8% Score 64; DB 1; Length 621;
Best Local Similarity 43.2% Pred. NO. 3.7;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 2 PKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTE 38
DB 175 PIVQIKGSGYPGEMCAGNELFTAGEDATISCVTE 211

RESULT 2
T41249
DEAD box ATP-dependent RNA helicase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: T41249
R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A:Reference number: 221981
A:Accession: T41249
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-604 <SEB>
A:Cross-references: EMBL:AL031545; PIDN:CAA20842.1; GSPDB:GN00066; SPDB:SPCC285.03
A:Experimental source: strain 972h-; cosmid c285
C:Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:41 ; Search time 2.40117 Seconds
(Without alignments) 690.935 Million cell updates/sec

Title: US-09-991-681-30
Sequence: 1 SPKVEKKDPSRKEMWENAGNKIYTAADKISKLTETK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	59.5	27.7	388 1	FTSM_HELPJ
2	58.5	27.2	104 1	SR19_ARCFU
3	56.5	26.3	388 1	FTSM_HELPJ
4	56	26.0	450 1	SR54_MYCPN
5	54	25.1	321 1	CS37_CANAL
6	54	25.1	387 1	YGH5_YEAST
7	53.5	24.9	466 1	PPCK_DEIRA
8	53.5	24.9	884 1	SECA_PORPU
9	53	24.7	343 1	Y130_METJA
10	53	24.7	446 1	SR54_MYCCE
11	52	24.2	840 1	SVL_BORBU
12	52	24.0	1956 1	ATX1_PLAFA
13	51.5	24.0	322 1	PR16_ARATH
14	51.5	24.0	747 1	ORPB_HUMAN
15	51	23.7	180 1	MATY_BACSH
16	51	23.7	230 1	Y771_METJA
17	51	23.7	547 1	EX42_DROME
18	51	23.7	572 1	YX37_MYCPN
19	51	23.7	605 1	AOR_PYRPU
20	51	23.7	802 1	PAC_ARTVI
21	51	23.7	2660 1	YEEJ_ECO57
22	51	23.7	4543 1	LRL1_CHICK
23	51	23.7	4544 1	LRL1_HUMAN
24	50.5	23.5	445 1	AS3A_MOUSE
25	50.5	23.5	745 1	YJ00_YEAST
26	50	23.3	208 1	GTH1_ARATH
27	50	23.3	209 1	GTH1_ARATH
28	50	23.3	359 1	SUR6_FUGRU
29	50	23.3	505 1	VLI_HPVO7
30	50	23.3	600 1	SP08_YEAST
31	50	23.3	859 1	ENV_EIYV3
32	50	23.3	1070 1	PVDR_PLAVS
33	50	23.3	1118 1	UBP8_HUMAN

34	50	23.3	1409 1	AEX3_CAEEL	002626 caenorhabd
35	50	23.3	1469 1	DP27_CAEEL	P48996 caenorhabd
36	50	23.3	2358 1	YEEJ_ECOLI	P76347 escherichia
37	49.5	23.0	178 1	EPAB_CHICK	P21760 gallus gall
38	49.5	23.0	235 1	YOF9_CAEEL	009336 caenorhabd
39	49.5	23.0	310 1	PP12_TOBAC	004857 nicotiana t
40	49	22.8	368 1	RALI_HUMAN	P78406 homo sapien
41	49	22.8	411 1	YG41_YEAST	P42942 saccharomyc
42	49	22.8	427 1	G6NT_BOVIN	092180 bos taurus
43	49	22.8	988 1	EPB3_CHICK	007458 gallus gall
44	49	22.8	1002 1	EPB5_CHICK	007457 gallus gall
45	49	22.8	1004 1	EPB2_CHICK	P28693 gallus gall

ALIGNMENTS

RESULT 1	FTSM_HELPJ	STANDARD:	PRT:	388 AA.
AC	Q9ZJ48:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Probable cell division protein ftsw.			
GN	FTSM OR JHP1468			
OS	Helicobacter pylori J99 (Campylobacter pylori J99).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OX	NCBI_Taxid=85963;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-99120557; PubMed-9923682;			
RA	Aim R.A., Ling U.-S.L., Moir D.T., King B.L., Brown E.D., Dols P.C.,			
RA	Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,			
RA	Tummino P.J., Caruso A., Urie-Nickelsen M., Mills D.M., Ives C.,			
RA	Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,			
RA	Trust T.J.;			
RT	"Genomic sequence comparison of two unrelated isolates of the human			
RT	gastric pathogen Helicobacter pylori."			
RL	Nature 397:176-180(1999).			
CC	-1- FUNCTION: THIS IS A SEPTUM-PEPTIDOGLYCAN BIOSYNTHETIC PROTEIN,			
CC	INVOLVED IN CELL WALL FORMATION. PLAYS A ROLE IN THE STABILIZATION			
CC	OF THE FTSM RING DURING CELL DIVISION (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(by similarity).			
CC	-1- SIMILARITY: BELONGS TO THE FTSM/RODA/SPOVE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL: AEO01568; AAO07042.1; -			
DR	InterPro: IPR001182; Cell cycle.			
DR	Pfam: PF01098; FTSM_RODA_SPOVE. 1.			
DR	PROSITE: PS00428; FTSM_RODA_SPOVE. 1.			
KW	Peptidoglycan synthesis; Cell wall; Cell division; Cell shape;			
FT	Transmembrane; Inner membrane; Complete proteome.			
FT	TRANSMEM 7			
FT	TRANSMEM 39			
FT	TRANSMEM 51			
FT	TRANSMEM 71			
FT	TRANSMEM 91			
FT	TRANSMEM 106			
FT	TRANSMEM 126			
FT	TRANSMEM 145			
FT	TRANSMEM 165			
FT	TRANSMEM 177			
FT	TRANSMEM 197			
FT	TRANSMEM 262			
FT	TRANSMEM 282			
FT	TRANSMEM 295			
FT	TRANSMEM 315			
FT	TRANSMEM 328			
FT	TRANSMEM 348			
FT	TRANSMEM 359			
FT	TRANSMEM 379			
SO	SEQUENCE 388 AA; 42540 MW; 51E4277A910B286D CRC64;			

Query Match 27.7% Score 59.5; DB 1; Length 388;
 Best Local Similarity 37.1%; Pred. No. 3;
 Matches 13; Conservative 4; Mismatches 13; Indels 5; Gaps 1;

OY 11 RKKEWENAGNKITMAADK-----TISKLMTEK 40
 DB 217 RLKLMWSNLQNSFTLLPDKLANALRISDLPESTYO 251

RESULT 2

SR19_ARCEU STANDARD; PRT; 104 AA.
 AC 029010;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Signal recognition particle 19 kDa protein (SRP19).
 GN SRP19 OR AF1258.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]

RP SEQUENCE FROM N.A.
 RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervilave A.R., Graham D.E., Kyprides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Ullrich T.,
 RA Colton M.D., Spillings T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing, archaeon Archaeoglobus fulgidus";
 RL Nature 390:364-370(1997).
 RN [2]

RP CHARACTERIZATION.
 RX MEDLINE=20150251; PubMed=10684931;
 RA Bhuiyan S.H., Gowda K., Hotokezaka H., Zwiab C.;

RT "Assembly of archaeal signal recognition particle from recombinant
 RT components";
 RL Nucleic Acids Res. 28:1365-1373(2000).
 RN [3]

RP CHARACTERIZATION.
 RX MEDLINE=20496765; PubMed=11041851;
 RA Diener J.L., Wilson C.;

RT "Role of SRP19 in assembly of the Archaeoglobus fulgidus signal
 RT recognition particle";
 RL Biochemistry 39:12862-12874(2000).
 RN [4]

CC -1- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY, BINDS DIRECTLY
 CC TO 7S RNA AND MEDIATES BINDING OF THE 54 KDA SUBUNIT OF THE SRP.
 CC -1- SUBUNIT: ARCHAEOAL SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA
 CC MOLECULE OF 300 NUCLEOTIDES AND TWO PROTEIN SUBUNITS: SRP54 AND
 CC SRP19.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001017; AAB89988.1;
 CC TIGR; AF1258;

DR InterPro: IPR002778; SRP19.
 DR Pfam: PF01922; SRP19; 1.
 DR ProDom: PD006609; SRP19; 1.
 KW Signal recognition particle; RNA-binding; Ribonucleoprotein;
 KW Complete proteome.
 SQ SEQUENCE 104 AA; 12405 MW; 7205DDBAB4E9E64 CRC64;

Query Match 27.2% Score 58.5; DB 1; Length 104;
 Best Local Similarity 44.1%; Pred. No. 1;
 Matches 15; Conservative 1; Mismatches 13; Indels 5; Gaps 2;

OY 5 EKKPSRRKKEWENAGNKITMAADKITSKLMTE 38
 DB 52 EKKYP---KSWMEGGRVYVEKRGKT--KLIME 80

RESULT 3

FTSM_HELPY STANDARD; PRT; 388 AA.
 AC P36096;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cell division protein ftsm.
 GN FTSM OR HP1560.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]

RP SEQUENCE FROM N.A.
 RX STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kervilave A.R., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Kirkness E.F., Peterson S.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodex A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Ullrich T.R., Peterson J.D., Kelley J.M.,
 RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori";
 RL Nature 388:539-547(1997).
 RN [2]

RP INVOLVED IN CELL WALL FORMATION. PLAYS A ROLE IN THE STABILIZATION
 CC OF THE FTSM2 RING DURING CELL DIVISION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (BY SIMILARITY).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000654; AAD08600.1;
 CC TIGR; HP1560;

DR InterPro: IPR001182; Cell cycle.
 DR Pfam: PF01098; FTSM_R0DA_SPOVE; 1.
 DR ProSite: PS00428; FTSM_R0DA_SPOVE; 1.
 KW Peptidoglycan synthesis; Cell wall; Cell division; Cell shape;
 KW Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 7 27 POTENTIAL.
 FT TRANSMEM 39 59 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 145 165 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.

FT TRANSMEM 262 282 POTENTIAL.
 FT TRANSMEM 295 315 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 359 379 POTENTIAL.
 SQ SEQUENCE 388 AA; 42478 MW; 77195EA8052BBF1 CRC64;

Query Match
 Best Local Similarity 26.3%; Score 56.5; DB 1; Length 388;
 Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 1;

OY 11 RKKEMENAGNKITYTMAADK-----TISKMTETK 40
 DB 217 RLKIMWSNLQNSLFTLLPRLANALRISDEPESTQ 251

RESULT 4
 SR54_MYCPN STANDARD; PRT; 450 AA.
 ID SR54_MYCPN
 AC P75054;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Signal recognition particle protein (fifty-four homolog).
 GN ffp OR MP061 OR MP093.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreid R., Hilbert H., Plagens H., Pirkil E., Li B.-C.,
 Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- FUNCTION: NECESSARY FOR EFFICIENT EXPORT OF EXTRA-CYTOPLASMIC PROTEINS. BINDS TO THE SIGNAL SEQUENCE WHEN IT EMERGES FROM THE RIBOSOMES (BY SIMILARITY).
 CC -1- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A SMALL CYTOPLASMIC RNA (SC-RNA) MOLECULE AND PROTEIN FFP (BY SIMILARITY).
 CC -1- DOMAIN: THE PROTEIN HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN BINDS GTP; THE M-DOMAIN BINDS THE RNA AND ALSO BINDS THE SIGNAL SEQUENCE.
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE000011; AAB95741.1; -
 DR HSSP; 007347; 1FFH.
 DR InterPro: IPR000897; SRP54.
 DR InterPro: IPR004125; SRP54-SPB.
 DR InterPro: IPR004780; SRP-sub.
 DR Pfam; PF00448; SRP54; 1.
 DR Pfam; PF02881; SRP54-N; 1.
 DR Pfam; PF02978; SRP-SPB; 1.
 DR ProDom; PD000819; SRP54; 1.
 DR TRIGRAMS; TIGR00959; 3a0501s01; 1.
 DR PROSITE; PS00300; SRP54; 1.
 KW Signal recognition particle; GTP-binding; RNA-binding;
 KM Complete proteome.
 FT DOMAIN 1 293 G-DOMAIN.
 FT DOMAIN 294 450 M-DOMAIN.
 FT NP_BIND 106 113 GTP (BY SIMILARITY).
 FT NP_BIND 188 192 GTP (BY SIMILARITY).
 FT NP_BIND 246 249 GTP (BY SIMILARITY).
 SQ SEQUENCE 450 AA; 50141 MW; 7EEF42EECAFE679 CRC64;

Query Match
 Best Local Similarity 26.0%; Score 56; DB 1; Length 450;
 Matches 13; Conservative 9; Mismatches 10; Indels 6; Gaps 1;

OY 2 PKVEKKDPSRKKEMENAGNKITYTMAADKTISKLTETK 39
 DB 387 PKLNRDPSRKRRIKSGRKM-----DELTKMKEM 418

RESULT 5
 CS37_CANAL STANDARD; PRT; 321 AA.
 ID CS37_CANAL
 AC P53707;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 37 kDa cell surface protein.
 GN CSP37 OR CSP30.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 26555;
 RX MEDLINE=97386400; PubMed=9244249;
 RA Sentandreu M., Nieto A., Iborra A., Elorza M.V., Ponton J.,
 Fonzi W.A., Sentandreu R.;

RT "Cloning and characterization of CSP37, a novel gene encoding a putative membrane protein of Candida albicans."
 RL J. Bacteriol. 179:4654-4663(1997).
 RN [2]

RP SEQUENCE OF 188-321 FROM N.A.
 RC STRAIN=ATCC 26555;
 RX MEDLINE=95387219; PubMed=7658300;
 RA Sentandreu M., Elorza M.V., Valentín E., Sentandreu R., Gozalbo D.;
 RT "Cloning of cDNAs coding for Candida albicans cell surface proteins."
 RL J. Med. Vet. Mycol. 33:105-111(1995).
 CC -1- SUBCELLULAR LOCATION: Cell wall.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U89676; AAB6368.1; -
 DR EMBL; S79004; AAB35168.1; -
 KW Cell wall.
 SQ SEQUENCE 321 AA; 37072 MW; BCSFP4D7AE119D00 CRC64;

Query Match
 Best Local Similarity 25.1%; Score 54; DB 1; Length 321;
 Matches 13; Conservative 8; Mismatches 13; Indels 2; Gaps 1;

OY 1 SPKVEKKDPSRKKEMENAGNKITYTMA--ADKTISK 34
 DB 137 SPNVEKKEKISFGNMFDSKNDKVDLKNADKKINE 172

RESULT 6
 YGH5_YEAST STANDARD; PRT; 387 AA.
 ID YGH5_YEAST
 AC P53159;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Hypothetical 44.6 kDa protein in RPL7A-HSF1 Intergenic region.
 GN YGI075C.
 OS Saccharomyces cerevisiae (Baker's yeast).

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001950; AAF10554.1; ALT_INIT.
DR HSPB; P22259; IAYL.
DR TIGR; DR0977. -.
DR InterPro; IPR001272; PEPCK_ATP.
DR Pfam; PF01293; PEPCK_ATP; 1.
DR ProDom; PD004723; PEPCK_ATP; 1.
DR TIGRFAMS; TIGR00224; pckA; 1.
DR PROSITE; PS00532; PEPCK_ATP; 1.
KW Glucoseogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.
FT NP_BIND 237 244 ATP (BY SIMILARITY).
SQ SEQUENCE 466 AA; 50951 MW; 22FDC11A2321E942 CRC64;

OY 1 SPR---VEKKDPSRKKEWMEAGKKITTMADKISKLM 36
Db ||| : : : : : ||| : : : : :
62 SPKDRFVDDQTRKTVWSEGFNOPIAEEVFDRLDKMV 100

RESULT 8
SECA_PORPU SECA_PORPU STANDARD; PRT; 884 AA.
AC P51381.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Preprotein translocase seca subunit.
GN SECA.
OS Porphyra purpurea.
OC Chloroplast.
CC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munholland J.;
RT "Complete nucleotide sequence of the porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT AND IMPORT. SECA HAS A
CC CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF ATP TO THE TRANSFER OF
CC PRE-SECRETORY PROTEINS ACROSS THE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38804; AAC08267.1; -.
DR InterPro; IPR000185; SECA.
DR Pfam; PF01043; SECA_protein; 1.
DR PRINTS; PR00906; SECA.
DR TIGRFAMS; TIGR00963; seca; 1.
DR PROSITE; PS01312; SECA; 1.
KW Protein transport; ATP-binding; Chloroplast; Translocation; Transport.
FT NP_BIND 98 105 ATP (BY SIMILARITY).
SQ SEQUENCE 884 AA; 101325 MW; A415846D12B902B CRC64;

Query Match 24.9%; Score 53.5; DB 1; Length 884;
Best Local Similarity 28.1%; Pred. No. 39;
Matches 16; Conservative 9; Mismatches 7; Indels 25; Gaps 3;

```


OY 9 PSRRKEV-----MENAGNKIX-----TMADKT--ISKIMTEYK 40
 DB 394 PLRRKESDLYVSNRYKMEAIADCEYDMTRAGRPLTVGTTSYKSLSKLITREYK 450

RESULT 9

Y130_METJA STANDARD: PRT: 343 AA.

AC 057594;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M0130.

OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_Taxid=2190;

RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; Pubmed=688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RT Science 273:1058-1073(1996).

RESULT 10
 SR54_MYCGE STANDARD: PRT: 446 AA.
 ID SR54_MYCGE
 AC P47294;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Signal recognition particle protein (fifty-four homolog).

OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_Taxid=2097;

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; Pubmed=7569993;

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RA "The minimal gene complement of Mycoplasma genitalium.";
 RT Science 270:397-403(1995).

CC -1- FUNCTION: NECESSARY FOR EFFICIENT EXPORT OF EXTRA-CYTOPLASMIC
 PROTEINS. BINDS TO THE SIGNAL SEQUENCE WHEN IT EMERGES FROM THE
 RIBOSOMES (BY SIMILARITY).

CC -1- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A SMALL
 CYTOPLASMIC RNA (SC-RNA) MOLECULE AND PROTEIN FRH (BY SIMILARITY).

CC -1- DOMAIN: THE PROTEIN HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN
 BINDS GTP; THE M-DOMAIN BINDS THE RNA AND ALSO BINDS THE SIGNAL
 SEQUENCE.

CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: U03684; AAC71264.1; -;
 DR HSSP: O07347; IFFH.
 DR TIGR: MG048; -;
 DR InterPro: IPR000897; SRP54.
 DR InterPro: IPR004125; SRP54_SBP.
 DR InterPro: IPR004780; SRP_sub.
 DR Pfam: PF00448; SRP54_1.
 DR Pfam: PF02881; SRP54_N; 1.
 DR Pfam: PF02978; SRP_SBP; 1.
 DR Prodom: PD000819; SRP54_1.
 DR TIGRFAMs: TIGR00959; 340501s01; 1.
 DR PROSITE: PS00300; SRP54; 1.

KW Signal recognition particle; GTP-binding; RNA-binding;
 KW Complete proteome.
 FT DOMAIN 1 293 G-DOMAIN.
 FT DOMAIN 294 446 M-DOMAIN.
 FT NP_BIND 106 113 GTP (BY SIMILARITY).
 FT NP_BIND 188 192 GTP (BY SIMILARITY).
 FT NP_BIND 246 249 GTP (BY SIMILARITY).
 SQ SEQUENCE 446 AA; 50199 MW; 080F2255C8B5E5AE CRC64;

Query Match 24.7%; Score 53; DB 1; Length 446;
 Best Local Similarity 31.6%; Pred. No. 23;
 Matches 12; Conservative 10; Mismatches 10; Indels 6; Gaps 1;

OY 2 PKVEKDPSPRRKEMWENAGNKIYTMADKTSKIMTEYK 39
 DB 387 PKLINRDPNKRKRITKSGRKM-----DELKMLAKEN 418

RESULT 11

DE Serine/threonine protein phosphatase PPI isozyme 6 (EC 3.1.3.16).
GN TOPP6 OR PPI6G OR AT4G1240 OR FBL21.30.
OS Arabidopsis thaliana (Mouse-ear cress).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95291339; Pubmed=7773310;
RA Arundhati A., Feiler H., Traas J., Zhang H., Lunness P.A.,
RA Doonan J.H.;
RT "A novel Arabidopsis type 1 protein phosphatase is highly expressed
RT in male and female tissues and functionally complements a conditional
RT cell cycle mutant of *Aspergillus*.";
RL Plant J. 7:823-834(1995).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=98278376; Pubmed=9617814;
RA Lin Q., Li J., Smith R.D., Walker J.C.;
RT "Molecular cloning and chromosomal mapping of type one
RT serine/threonine protein phosphatases in *Arabidopsis thaliana*.";
RL Plant Mol. Biol. 37:471-481(1998).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; Pubmed=10617198;
RA Meyer K.F.X., Schueller C., Mambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duescherhoeft A., Stiekema W., Enlhan K.-D., Terryn N.,
RA Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Macho R., Mueller M.,
RA Kreis M., Delsen M., Pulidomench P., Watson M., Schmidtnehl T.,
RA Reichert B., Portellelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hobeisel J., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Gymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Breken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,
RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dikse W.,
RA Moolman P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,
RA Beneliser S., Hempel S., Feldpausch M., Lambrecht S., Van den Dele H.,
RA De Keyser A., Buysnaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
RA Petrett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argitlon A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muehlen A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Baynes M., Terol J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Hebermann K.,
RA Parnell L., Dedha N., Gnoj L., Schütz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kallicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mux P., Bentley D., Fulton L., Mardis E., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Dante B., Pepin K., Hillier L.,
RA Nelson J., Spiehl J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonciu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Grenat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martensen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.";
RL Nature 402:769-777(1999).

CC -{- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O - a protein +
CC phosphate.
CC -{- TISSUE SPECIFICITY: STRONGLY UPREGULATED WITHIN DEVELOPING
CC FLOWERS, ESPECIALLY IN THE TAPETUM, THE DEVELOPING AND MATURE
CC POLLEN AND IN THE OVARIES.
CC -{- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z46253; CAAB6339.1; -;
CC EMBL; U80921; AAC39460.1; -;
CC EMBL; AL096882; CAB51408.1; -;
CC EMBL; AL161531; CAB81225.1; -;
CC HSSP; P08129; IRFM;
CC InterPro; IPR004843; M-peptidase.
CC InterPro; IPR004844; S/T-phosphatase.
CC Pfam; PF00149; Metallophos; 1.
CC PRINTS; PR00114; STPHPTASE.
CC ProDom; PD000252; S/T-phosphatase; 1.
CC SMART; SM00156; PP2Ac; 1.
CC PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
CC HYDROLASE; Iron; Manganese; Multigene family.
CC METAL 61 61
CC METAL 63 63 IRON (BY SIMILARITY).
CC METAL 89 89 IRON (BY SIMILARITY).
CC METAL 121 121 IRON AND MANGANESE (BY SIMILARITY).
CC METAL 122 122 MANGANESE (BY SIMILARITY).
CC ACT_SITE 170 170 GENERAL ACID (BY SIMILARITY).
CC METAL 170 170 MANGANESE (BY SIMILARITY).
CC METAL 245 245 MANGANESE (BY SIMILARITY).
CC CONFLICT 248 248 D -> V (IN REF. 3).
CC CONFLICT 318 322 GKGR -> V (IN REF. 3).
CC SEQUENCE 322 AA; 36582 MW; B23218743E66205 CRC64;

Oy 8 DPSRK-KEMMENACNKTITYNADTKISKLMT 37
Db 205 DPKREIQGKENGKGVSTTGADKVAERLQT 235

Query Match 24.0% Score 51.5; DB 1; Length 322;
Best Local Similarity 41.9% Pred. No. 25;
Matches 13; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

RESULT 14
ORPB_HUMAN
ID ORPB_HUMAN STANDARD; PRT; 747 AA.
AC Q9BXB4; STANDARD; PRT; 747 AA.
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxyterol binding protein-related protein 11 (OSBP-related protein 11)
DE (OSBP-11).
GN OSBP11 OR ORP11 OR OSBP12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21601154; Pubmed=11735225;
RA Javorksi C.J., Moreira E., Li A., Lee R., Rodriguez I.R.;
RT "A family of 12 human genes containing oxyterol-binding domains.";
RL Genomics 78:185-196(2001).
RN [2]
RN SEQUENCE OF 1-315 FROM N.A.
RX MEDLINE=21376257; Pubmed=11483621;
RA Lehto M., Laitinen S., Chinetti G., Johansson M., Ehnholm C.,

```

RA   Staelens B., Ikonen E., Olkkonen V.M.:
RT   "The OSBP-related protein family in humans.";
RL   J. Lipid Res. 42:1203-1213(2001).
CC   -1- TISSUE SPECIFICITY: Widely expressed.
CC   -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC   -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC   or send an email to license@isb-slb.ch).
CC   -----
CC   EMBL; AF392454; AAL40667.1; -
CC   EMBL; AF346292; AAK31141.1; -
CC   Genew; HGNC:16397; OSBPL1L.
CC   MIM: 606739; -
DR   InterPro; IPR000648; Oxysterol_BP.
DR   InterPro; IPR01849; PH.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF01237; Oxysterol_BP; 1.
DR   SMART; SM00233; PH; 1.
DR   PROSITE; PS01013; OSBP; FALSE_NEG.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   Lipid transport; Transport.
DR   T_T DOMAIN 58 155 PH.
SO   SEQUENCE 747 AA; 83643 MW; B443D3BDE8A5F6B CRC64;

Query Match 24.0%; Score 51.5; DB 1; Length 747;
Best Local Similarity 29.7%; Pred. NO. 59;
Matches 11; Conservative 7; Mismatches 14; Indels 5; Gaps 1;

0Y 4 VEKKDPSRKKEWMENAGNKIITYTMADRTISKIMLEYK 40
   :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 666 LEKODPFESRILMKNVTDLSRESEIDIA-----TEHK 697

```

```

RESULT 15
MATA_BACSH
ID MATA_BACSH STANDARD: PRI: 180 AA.
AC P26840.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Probable macrolide acetyltransferase (EC 2.3.1.-) (Fragment).
OS Bacillus sphaericus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1421;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=87083389; PubMed=3025178;
RA Monod M., Mohan S., Dubnau D. :
RT "Cloning and analysis of ermC, a new
RT macrolide-lincosamide-streptogramin B resistance element from
RT Bacillus sphaericus." ;
RL J. Bacteriol. 169:340-350(1987).
CC -1- SIMILARITY: BELONGS TO THE CSE/LACA/LDPA/MODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15332; AAA22417.1; ALT_INT.
CC HSSP: P26841; 2XAT.
CC InterPro: IPR001451; Hexapep_transf.
CC Pfam: PF00132; hexapep; 3.

```

DR PROSITE: PS00101; HEXAPEP_TRANSFERRASE; 1.
KW Antibiotic resistance; Transferase; Acyltransferase; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 180 AA; 20399 MW; 7EDF56157E6FB1A CRC64;

Query Match 23.7%; Score 51; DB 1; Length 180;
Best Local Similarity 28.2%; Pred. NO. 16;
Matches 11; Conservative 10; Mismatches 14; Indels 4; Gaps 1;

QY 5 EKKDPSRRKKEWENAGNKIV----TMAADTKISKIMLEY 39
||| ||| : : : : :
Db 135 EKIEFLLEKLEWMNNSGEEIFPNLELITISSEALDELMNNY 173

Search completed: June 17, 2003, 11:56:48
Job time : 3.40117 secs

PA (ROBE/) ROBERTS-RAPP L.
 XX Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-187683/24.
 DR
 XX Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 PS Claim 17; Page 44; 57pp; English.
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 283-317 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 180; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1a-18; 0;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
 DB 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
 RESULT 2
 AAM85472
 ID AAM85472 standard; Protein; 518 AA.
 XX
 AC AAM85472;
 XX
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN MO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98MO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
 PI Robert-Strapp L, Russell JC, Stroupe SD;

XX WPI; 1998-610000/51;
 DR N-PSDB; AAV82812.
 XX
 PT New P118 nucleic acid and proteins - used for diagnosis and
 PT treatment of prostatic disease, especially cancer, and also for drug
 PT screening
 XX
 PS Claim 17; Page 93-94; 117pp; English.
 CC The present sequence is encoded by consensus PS118 sequence derived from
 CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
 CC were identified from a cDNA library made from prostate tumour tissue.
 CC Recombinant PS118 protein is used to detect PS118-specific antibodies,
 CC to raise antibodies for detection of PS118 antigens, to screen for
 CC specific binding agents (potential therapeutics), and to isolate specific
 CC antibodies from serum. Detection of PS118 protein or nucleic acid, which
 CC are prostate related, and altered or elevated in prostatic disease, is
 CC used for detection, diagnosis, staging, monitoring and prognosis of
 CC prostatic disease, particularly cancer, and to identify subjects at
 CC risk.
 SQ Sequence 518 AA;
 Query Match 100.0%; Score 180; DB 19; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2,4e-17;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 317
 DB 283 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 317
 RESULT 3
 AAM50809
 ID AAM50809 standard; Protein; 518 AA.
 XX
 AC AAM50809;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker partial sequence.
 XX
 DE PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTs T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLASs M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-187683/24.

DR N-PSDB; ABA91651.
XX
PT Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
PS
PS Claim 17; Page 42-43; 57pp; English.
XX
CC The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus cDNA sequence
CC (see ABA91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia and conditions of the
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.
XX
SQ Sequence 518 AA;
XX
Query Match 100.0%; Score 180; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.4e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFQSESTPSTGFGSKETPSDDRSQSRHMGES 35
DB 283 SFQSESTPSTGFGSKETPSDDRSQSRHMGES 317
RESULT 4
ABG09728
ID ABG09728 standard; Protein: 1807 AA.
XX
AC ABG09728;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9719.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS73915.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
PS Claim 20; SEQ ID NO 40087; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1807 AA;
XX
Query Match 100.0%; Score 180; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFQSESTPSTGFGSKETPSDDRSQSRHMGES 35
DB 1572 SFQSESTPSTGFGSKETPSDDRSQSRHMGES 1606
RESULT 5
ABG09731
ID ABG09731 standard; Protein: 1982 AA.
XX
AC ABG09731;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9722.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS73918.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 253 AA:

Query Match 32.8%; Score 59; DB 22; Length 253;
Best Local Similarity 42.9%; Pred. No. 2.7;
Matches 15; Conservative 5; Mismatches 11; Indels 4; Gaps 2;

OY 4 SESSTST---GFSGK-ETPSDDRSQREHMG 34
Db 122 SATSTPSYLPGLGDKSETPSEEEEEEERGE 156
| : | | | | : | : | | | | | : | | |

RESULT 8
ABB60540

ID ABB60540 standard; Protein; 1403 AA.

AC ABB60540;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 8412.

XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.

XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75;

DR N-PDB; ABL04643.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Disclosure; SEQ ID NO 8412; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC sequences (ABU1840-ABU16175) and the encoded proteins
CC (ABBS7731-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 1403 AA:

Query Match 32.8%; Score 59; DB 22; Length 1403;
Best Local Similarity 38.1%; Pred. No. 20;

Matches 16; Conservative 5; Mismatches 13; Indels 8; Gaps 1;

OY 1 SFSSESSTSTGFSGKRT-----PSDDRSQREHMG 34
Db 1104 SSSSSSTSLSLAGEESDFNELALEXELSSQGEHMG 1145
| : | | | | : | : | | | | | : | | |

RESULT 9

AAU42806

ID AAU42806 standard; Protein; 73 AA.

AC AAU42806;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #3702.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertostis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS WO200181581-A2.

PN 01-NOV-2001.

PF 20-APR-2001; 2001MO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

DR L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR N-PDB; AAS59518.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

PS Example 1; SEQ ID NO 4001; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertostis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 73 AA:

Query Match 32.2%; Score 58; DB 22; Length 73;
Best Local Similarity 34.3%; Pred. No. 0.88;

XX DE Protein #2602 encoded by probe for measuring heart cell gene expression.
 XX XX
 KM Human: gene expression: heart; microarray: vascular system;
 KM cardiovascular disease; hypertension; cardiac arrhythmia;
 KM congenital heart disease.
 OS Homo sapiens.
 XX PN WO200157274-A2.
 XX XX
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00666.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488899/53.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX PS Claim 15; SEQ ID No 22373; 530bp; English.
 XX SQ Sequence 102 AA;
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC AB21535-AB41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX
 SQ Query Match 31.1%; Score 56; DB 22; Length 102;
 Best Local Similarity 40.0%; Pred. No. 2.5;
 Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;
 QY 3 QSESSTPSTGFGSKETPSEDDRSQSRHM 32
 Db 27 QNESQSPQ---EPEEGPSEDDKAEGEEM 52
 RESULT 13
 AAM56001
 ID AAM56001 standard; Protein; 102 AA.
 XX AC AAM56001;
 XX XX
 DT 05-NOV-2001 (first entry)
 XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28106.
 XX KM Human: brain expressed exon; gene expression analysis; probe;
 KM microarray: Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 OS Homo sapiens.

XX PN WO200157275-A2.
 XX XX
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00667.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-483446/52.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX PS Example 4; SEQ ID NO: 28106; 650bp + Sequence Listing; English.
 XX SQ Sequence 102 AA;
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 CC XX
 SQ Query Match 31.1%; Score 56; DB 22; Length 102;
 Best Local Similarity 40.0%; Pred. No. 2.5;
 Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;
 QY 3 QSESSTPSTGFGSKETPSEDDRSQSRHM 32
 Db 27 QNESQSPQ---EPEEGPSEDDKAEGEEM 52
 RESULT 14
 AAM68370
 ID AAM68370 standard; Protein; 102 AA.
 XX AC AAM68370;
 XX XX
 DT 06-NOV-2001 (first entry)
 XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28676.
 XX KM Human: bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; Leukaemia; lymphoma; myeloma.
 OS Homo sapiens.
 XX PN WO200157276-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00668.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:56:11 ; Search time 3.2284 Seconds

(without alignments)
318.982 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFQSESTPTGTGFGSKETPESDDRSQSRHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCYUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfill1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	100.0	35	US-09-065-383-29	Sequence 29, Appl
2	180	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	52.5	29.2	717	US-08-910-925-1	Sequence 1, Appl
4	50	27.8	309	US-08-465-167A-24	Sequence 24, Appl
5	50	27.8	309	US-08-993-118-10	Sequence 10, Appl
6	50	27.8	309	US-08-845-528C-10	Sequence 10, Appl
7	50	27.8	309	US-08-627-820-24	Sequence 24, Appl
8	50	27.8	401	US-08-549-004A-5	Sequence 5, Appl
9	50	27.8	401	US-09-051-982A-5	Sequence 5, Appl
10	50	27.8	1261	US-09-208-742-4	Sequence 4, Appl
11	50	27.8	1261	US-09-332-295-2	Sequence 2, Appl
12	50	27.8	1261	US-09-709-979-2	Sequence 2, Appl
13	49.5	27.5	674	US-08-893-852A-1	Sequence 1, Appl
14	49.5	27.5	703	US-08-910-925-4	Sequence 4, Appl
15	49	27.2	683	US-09-620-412C-357	Sequence 357, App
16	49	27.2	821	US-09-556-877-195	Sequence 195, App
17	49	27.2	821	US-09-620-412C-195	Sequence 195, App
18	49	27.2	1776	US-09-556-877-179	Sequence 179, App
19	49	27.2	1776	US-09-620-412C-179	Sequence 179, App
20	48.5	26.9	568	US-08-320-559-30	Sequence 30, Appl
21	48.5	26.9	568	US-08-545-860D-30	Sequence 30, Appl
22	48.5	26.9	568	PCT-US94-04496-30	Sequence 30, Appl
23	48	26.7	714	US-09-347-878-16	Sequence 16, Appl
24	47.5	26.4	590	US-08-893-852A-4	Sequence 4, Appl
25	47.5	26.4	590	US-08-821-818-2	Sequence 2, Appl
26	47	26.1	300	US-09-277-078-2	Sequence 2, Appl
27	47	26.1	365	US-08-481-814A-10	Sequence 10, Appl

28	47	26.1	1704	US-08-485-355B-40	Sequence 40, Appl
29	47	26.1	2265	US-08-149-097D-36	Sequence 36, Appl
30	47	26.1	2509	US-08-149-097D-35	Sequence 35, Appl
31	46.5	25.8	98	US-08-481-658B-50	Sequence 50, Appl
32	46.5	25.8	98	US-08-477-504A-50	Sequence 50, Appl
33	46.5	25.8	98	US-08-486-756A-50	Sequence 50, Appl
34	46.5	25.8	98	US-08-485-862B-50	Sequence 50, Appl
35	46.5	25.8	98	US-08-487-077A-50	Sequence 50, Appl
36	46.5	25.8	98	US-08-485-863A-50	Sequence 50, Appl
37	46.5	25.8	106	US-08-485-049D-50	Sequence 50, Appl
38	46.5	25.8	240	US-08-893-042-1	Sequence 1, Appl
39	46.5	25.8	240	US-08-760-745-3	Sequence 3, Appl
40	46.5	25.8	377	US-08-787-739-87	Sequence 87, Appl
41	46.5	25.8	377	US-09-178-115-87	Sequence 87, Appl
42	46.5	25.8	377	US-09-177-776-87	Sequence 87, Appl
43	46.5	25.8	422	US-08-335-469-2	Sequence 2, Appl
44	46.5	25.8	459	US-08-481-658B-2	Sequence 2, Appl
45	46.5	25.8	459	US-08-477-504A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANDOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON R.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084-US-P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-29

Query Match 100.0%; Score 180; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFOSESTPSTGSGKTEPSEDDRSQSRHMGES 35
Db 1 SFOSESTPSTGSGKTEPSEDDRSQSRHMGES 35

RESULT 2

US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPF, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084,US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ. ID NO. 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 180; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SFOSESTPSTGSGKTEPSEDDRSQSRHMGES 35
Db 283 SFOSESTPSTGSGKTEPSEDDRSQSRHMGES 317

RESULT 3

US-08-910-925-1
Sequence 1, Application US/08910925
Patent No. 6162601

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Puri
TITLE OF INVENTION: HUMAN PININ SPLICER VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,925
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0365 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 717 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: FIBROTO1
CLONE: 53219
US-08-910-925-1

Query Match 29.2%; Score 52.5; DB 4; Length 717;
Best Local Similarity 41.2%; Pred. No. 27;
Matches 14; Conservative 6; Mismatches 9; Indels 5; Gaps 1;

QY 1 SFOSESTPSTGSGKTEPSEDDRSQSR 29
Db 603 SRSSSSSSSTGGSSSRDSSSTSSSSSESRSR 636

RESULT 4

US-08-465-167A-24
Sequence 24, Application US/08465167A
Patent No. 5750395

GENERAL INFORMATION:

APPLICANT: Fikes, John D.
APPLICANT: Livingston, Brian D.
APPLICANT: Sette, Alessandro D.
APPLICANT: Sidney, John C.

```

      FILING DATE:
      CLASSIFICATION:
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/845,528
      FILING DATE: April 25, 1997
      ATTORNEY/AGENT INFORMATION:
      NAME: Mary Anne Schofield
      REGISTRATION NUMBER: 36,669
      REFERENCE/DOCKET NUMBER: LUD 5455
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 688-9200
      TELEFAX: (212) 838-3884
      INFORMATION FOR SEQ ID NO: 10:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 309
      TYPE: amino acids
      STRANDEDNESS: single stranded
      TOPOLOGY: linear
      US-08-993-118-10

Query Match          27.8%; Score 50; DB 2; Length 309;
Best Local Similarity 42.4%; Pred. No. 22;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY      3 QSESSRPSGTGFGSGKETPSEDDRSQSRHMGES 35
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      59 QGASAPFTYINFTROQRPSEG--SSSRREEGSPS 89

RESULT 6
US-08-845-528C-10
: Sequence 10, Application US/0845528C
: Patent No. 6027924
: GENERAL INFORMATION:
: APPLICANT: LUCAS, Sophie;
: APPLICANT: DE SMET, Charles;
: TITLE OF INVENTION: BOON-FALBEUR, Thierry
: TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
: TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felife & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/845,528C
: FILING DATE: April 25, 1997
: CLASSIFICATION: 4335
: ATTORNEY/AGENT INFORMATION:
: NAME: Mary Anne Schofield
: REGISTRATION NUMBER: 36,669
: REFERENCE/DOCKET NUMBER: LUD 5455
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200
: TELEFAX: (212) 838-3884
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 309
: TYPE: amino acids
: STRANDEDNESS: single stranded
: TOPOLOGY: linear
: US-08-845-528C-10

Query Match          27.8%; Score 50; DB 3; Length 309;

```


TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TMLR3DT01
CLONE: 508302
US-08-893-852A-1

Query Match 27.5% Score 49.5; DB 3; Length 674;
Best Local Similarity 37.9% Pred. No. 65;
Matches 11; Conservative 6; Mismatches 5; Indels 7; Gaps 1;

OY 3 QSESSTPSTGTF-----SGKEPSEDD 24
DB 374 EASSTPATGFLKSWYQPEDEEED 402

RESULT 14
US-08-910-925-4
Sequence 4, Application US/08910925
Patent No. 6162601
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Puri
TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910.925
FILING DATE: Herevalch
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0365 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 703 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1684843
US-08-910-925-4

Query Match 27.5% Score 49.5; DB 4; Length 703;
Best Local Similarity 38.2% Pred. No. 68;
Matches 13; Conservative 7; Mismatches 9; Indels 5; Gaps 1;

OY 1 SFQESSTPSTGTFSGKE-----TPSEDDRSQSR 29
DB 615 SSRSSSSSTSGSSSRDSSSSSTSSSRSSRSR 648

RESULT 15

US-09-620-412C-357
Sequence 357, Application US/09620412C
Patent No. 6448234.
GENERAL INFORMATION:
APPLICANT: Steven P. Filing
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C7
CURRENT APPLICATION NUMBER: US/09/620.412C
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FASTSEQ for Windows Version 3.0/4.0
SEQ ID NO 357
LENGTH: 683
TYPE: PRT
ORGANISM: Chlamydia trachomatis
US-09-620-412C-357

Query Match 27.2% Score 49; DB 4; Length 683;
Best Local Similarity 34.3% Pred. No. 78;
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

OY 1 SFQESSTPSTGTFSGKEPSEDDRSQSRHMGES 35
DB 173 SSRSSSTPDGGAASSGAPSGDQISANACLAKS 207

Search completed: June 17, 2003, 12:03:39
Job time : 4.2284 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:36 ; Search time 3.74085 Seconds
(Without alignments)
1000.321 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFQSESTPSTGCGSKETPSDDRSQREHMGFS 35

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	31.1	102	10 US-09-864-761-35901	Sequence 35901, A
2	53.5	29.7	737	9 US-10-228-931-4	Sequence 4, Appl1
3	52.5	29.2	91	9 US-10-090-035-16	Sequence 16, Appl1
4	51	28.3	42	10 US-09-955-807-9	Sequence 9, Appl1
5	51	28.3	65	10 US-09-955-807-5	Sequence 5, Appl1
6	51	28.3	65	10 US-09-955-807-10	Sequence 10, Appl1
7	51	28.3	77	10 US-09-955-807-4	Sequence 4, Appl1
8	51	28.3	79	10 US-09-955-807-3	Sequence 3, Appl1
9	51	28.3	80	10 US-09-955-807-16	Sequence 16, Appl1
10	51	28.3	89	10 US-09-955-807-15	Sequence 15, Appl1
11	51	28.3	105	10 US-09-955-807-2	Sequence 2, Appl1
12	51	28.3	105	10 US-09-955-807-14	Sequence 14, Appl1
13	51	28.3	758	10 US-09-904-887-5	Sequence 5, Appl1
14	50	27.8	212	9 US-09-764-891-4686	Sequence 4686, Ap
15	50	27.8	253	10 US-09-864-761-37733	Sequence 37733, A
16	50	27.8	275	10 US-09-755-456-9	Sequence 9, Appl1
17	50	27.8	309	9 US-09-766-889A-2	Sequence 2, Appl1
18	50	27.8	309	9 US-10-085-108-10	Sequence 10, Appl1
19	50	27.8	413	10 US-09-801-574-2	Sequence 2, Appl1

20	50	27.8	856	9 US-09-738-626-3515	Sequence 3515, Ap
21	50	27.8	1031	10 US-09-815-242-10932	Sequence 10932, A
22	50	27.8	1261	12 US-10-147-268-2	Sequence 2, Appl1
23	50	27.8	1770	10 US-09-841-132-444	Sequence 444, Appl
24	49.5	27.5	99	10 US-09-216-393-20	Sequence 20, Appl
25	49.5	27.5	515	10 US-09-801-368-384	Sequence 384, Appl
26	49.5	27.5	515	10 US-09-771-161A-195	Sequence 195, App
27	49	27.2	331	10 US-09-815-242-5912	Sequence 5912, Ap
28	49	27.2	331	10 US-09-815-242-131872	Sequence 12872, A
29	49	27.2	331	10 US-09-815-242-13146	Sequence 13146, A
30	49	27.2	489	10 US-09-876-889-350	Sequence 350, App
31	49	27.2	683	10 US-09-841-132-357	Sequence 357, App
32	49	27.2	821	10 US-09-841-132-195	Sequence 195, App
33	49	27.2	1776	10 US-09-841-132-179	Sequence 179, App
34	49	27.2	2011	9 US-10-176-847-56	Sequence 56, Appl
35	48.5	26.9	724	10 US-09-925-300-1053	Sequence 1053, Ap
36	48	26.7	121	9 US-09-764-891-4135	Sequence 4135, Ap
37	48	26.7	444	9 US-09-738-626-5577	Sequence 6577, Ap
38	48	26.7	971	9 US-10-108-605-317	Sequence 317, App
39	47.5	26.4	118	9 US-10-026-007-32	Sequence 32, Appl
40	47.5	26.4	435	9 US-10-000-512-18	Sequence 18, Appl
41	47.5	26.4	590	10 US-09-052-753-2	Sequence 2, Appl1
42	47.5	26.4	713	9 US-10-028-332-35	Sequence 35, Appl1
43	47.5	26.4	713	10 US-09-822-687-2	Sequence 2, Appl1
44	47	26.1	63	10 US-09-864-761-42311	Sequence 42311, A
45	47	26.1	266	10 US-09-864-761-37015	Sequence 37015, A

ALIGNMENTS

RESULT 1
US-09-864-761-35901
Sequence 35901, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 35901
;; LENGTH: 102
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL096678.8
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
;; OTHER INFORMATION: EST_HUMAN HIT: BF792924.1, EVALUO 1.40e-01
US-09-864-761-35901
```

```
Query Match          31.1%; Score 56; DB 10; Length 102;
Best Local Similarity 40.0%; Pred. No. 3;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;
```

```
OY 3 QSESSTPGFGSGKETPSDDRSQSRHM 32
Db 27 QNESQSPQ-----EPPEGSPEDDKAEAGEEM 52
```

```
RESULT 2
US-10-228-931-4
;; Sequence 4, Application US/10228931
;; Publication No. US20030051258A1
;; GENERAL INFORMATION:
;; APPLICANT: Verma, Ajit K
;; APPLICANT: Reddy, Peter J
;; APPLICANT: Jansen, Aaron P
;; TITLE OF INVENTION: Animal Model System for Squamous Cell Carcinoma
;; FILE REFERENCE: 960296.97613
;; CURRENT APPLICATION NUMBER: US/10/228,931
;; CURRENT FILING DATE: 2002-08-27
;; PRIOR APPLICATION NUMBER: US/09/772,647
;; PRIOR FILING DATE: 2001-01-30
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 737
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: T7 tag and
;; OTHER INFORMATION: mouse protein kinase C epsilon coding sequence
US-10-228-931-4
```

```
Query Match          29.7%; Score 53.5; DB 9; Length 737;
Best Local Similarity 48.0%; Pred. No. 60;
Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;
```

```
OY 4 SESSTPGFGSGKETPSDDRSQS 28
Db 327 AESPPASG-----NSPSEDRSKS 346
```

```
RESULT 3
US-10-090-035-16
```

```
;; Sequence 16, Application US/10090035
;; Patent No. US20020170089A1
;; GENERAL INFORMATION:
;; APPLICANT: Simmons, Carl R.
;; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
;; TITLE OF INVENTION: Proteins and Uses Thereof
;; FILE REFERENCE: 35718/242990
;; CURRENT APPLICATION NUMBER: US/10/090,035
;; CURRENT FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: 60/272,227
;; PRIOR FILING DATE: 02/28/2001
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 16
;; LENGTH: 91
;; TYPE: PRT
;; ORGANISM: Oryza sativa
US-10-090-035-16
```

```
Query Match          29.2%; Score 52.5; DB 9; Length 91;
Best Local Similarity 43.9%; Pred. No. 7.6;
Matches 18; Conservative 1; Mismatches 11; Indels 11; Gaps 2;
```

```
OY 1 SPQSESTPGFGSG-----KETPSDDRSQSRH 31
Db 11 SEEVRSYTP-TGGFLRGVQOQHVYKETFQETDRSGSH 50
```

```
RESULT 4
US-09-955-807-9
;; Sequence 9, Application US/09955807
;; Patent No. US20020132966A1
;; GENERAL INFORMATION:
;; APPLICANT: Lok, SI
;; APPLICANT: Sheppard, Paul O.
;; APPLICANT: Kindsvogel, Wayne
;; APPLICANT: Bort, Susan J.
;; TITLE OF INVENTION: Secretory Protein-48
;; FILE REFERENCE: 98-17C1
;; CURRENT APPLICATION NUMBER: US/09/955,807
;; CURRENT FILING DATE: 2001-09-19
;; PRIOR APPLICATION NUMBER: 60/102,679
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 09/410,603
;; PRIOR FILING DATE: 1999-10-01
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 9
;; LENGTH: 42
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-955-807-9
```

```
Query Match          28.3%; Score 51; DB 10; Length 42;
Best Local Similarity 52.2%; Pred. No. 5;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;
```

```
OY 9 PSTGEGSG-----KETPSDDRSQ 27
Db 17 PAVVGTGDSGAKETVSQDKRSQ 39
```

```
RESULT 5
US-09-955-807-5
;; Sequence 5, Application US/09955807
;; Patent No. US20020132966A1
;; GENERAL INFORMATION:
;; APPLICANT: Lok, SI
;; APPLICANT: Sheppard, Paul O.
;; APPLICANT: Kindsvogel, Wayne
;; APPLICANT: Bort, Susan J.
;; TITLE OF INVENTION: Secretory Protein-48
;; FILE REFERENCE: 98-17C1
```

```

; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 65
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-09-955-807-5

Query Match
Best Local Similarity 28.3%; Score 51; DB 10; Length 65;
Pred. No. 8.2;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 9 PSTGFGSG----KETPSEDDRSQ 27
   1: 11:1 111 1:1 111
Db 14 PAVTGFTGDSGAKETVSQDKRSQ 36

RESULT 6
US-09-955-807-10
; Sequence 10, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, SI
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 65
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-09-955-807-10

Query Match
Best Local Similarity 28.3%; Score 51; DB 10; Length 65;
Pred. No. 8.2;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 9 PSTGFGSG----KETPSEDDRSQ 27
   1: 11:1 111 1:1 111
Db 17 PAVTGFTGDSGAKETVSQDKRSQ 39

RESULT 7
US-09-955-807-4
; Sequence 4, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, SI
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 80
```

```

; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 77
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-09-955-807-4

Query Match
Best Local Similarity 28.3%; Score 51; DB 10; Length 77;
Pred. No. 10;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 9 PSTGFGSG----KETPSEDDRSQ 27
   1: 11:1 111 1:1 111
Db 26 PAVTGFTGDSGAKETVSQDKRSQ 48

RESULT 8
US-09-955-807-3
; Sequence 3, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, SI
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 79
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-09-955-807-3

Query Match
Best Local Similarity 28.3%; Score 51; DB 10; Length 79;
Pred. No. 10;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

OY 9 PSTGFGSG----KETPSEDDRSQ 27
   1: 11:1 111 1:1 111
Db 28 PAVTGFTGDSGAKETVSQDKRSQ 50

RESULT 9
US-09-955-807-16
; Sequence 16, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, SI
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 80
```

TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-16

Query Match
Best Local Similarity 52.2%; Score 51; DB 10; Length 80;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 9 PSTGFGSG---KETPSEDDRSQ 27
1: ||:| ||| |:
Db 29 PAVTGFTGDSGAKETVSQDKRSQ 51

RESULT 10
US-09-955-807-15
Sequence 15, Application US/099555807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955,807
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/102,679
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 09/410,603
PRIOR FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-15

Query Match
Best Local Similarity 52.2%; Score 51; DB 10; Length 89;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 9 PSTGFGSG---KETPSEDDRSQ 27
1: ||:| ||| |:
Db 38 PAVTGFTGDSGAKETVSQDKRSQ 60

RESULT 11
US-09-955-807-2
Sequence 2, Application US/099555807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955,807
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/102,679
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 09/410,603
PRIOR FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-2

Query Match 28.3%; Score 51; DB 10; Length 105;

Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 9 PSTGFGSG---KETPSEDDRSQ 27
1: ||:| ||| |:
Db 54 PAVTGFTGDSGAKETVSQDKRSQ 76

RESULT 12
US-09-955-807-14
Sequence 14, Application US/099555807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955,807
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/102,679
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 09/410,603
PRIOR FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-14

Query Match
Best Local Similarity 52.2%; Score 51; DB 10; Length 105;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 9 PSTGFGSG---KETPSEDDRSQ 27
1: ||:| ||| |:
Db 54 PAVTGFTGDSGAKETVSQDKRSQ 76

RESULT 13
US-09-904-987-5
Sequence 5, Application US/09904987
Patent No. US20020037908A1
GENERAL INFORMATION:
APPLICANT: No. US20020037908A1actyl, Inc.
TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
FILE REFERENCE: 42108/26146
CURRENT APPLICATION NUMBER: US/09/904,987
CURRENT FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 758
TYPE: PRT
ORGANISM: homo sapiens
PUBLICATION INFORMATION:
DATABASE ENTRY DATE: 2001-02-13
RELEVANT RESIDUES: (1)..(758)
US-09-904-987-5

Query Match
Best Local Similarity 27.3%; Score 51; DB 10; Length 758;
Matches 12; Conservative 7; Mismatches 13; Indels 12; Gaps 1;

QY 4 SESSTPSTGFGSGKETPSE-----DDRSQSGREHGES 35
1: ||:| ||| |:
Db 282 SEPDGFSVGRKQGDAPLEFTFVETPNVQKQASHEHLGRA 325

RESULT 14
US-09-764-891-4686
Sequence 4686, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4686
LENGTH: 212
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (62)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (197)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (205)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4686

Query Match 27.8%; Score 50; DB 9; Length 212;
Best Local Similarity 35.7%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 SFOSESSTPGTSGFSGKETPSEDDRSQS 28
Db 161 NLSSNLFSPQCGGLTGCKEAKKEDGTST 188

RESULT 15
US-09-864-761-37733
Sequence 37733, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37733
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006059.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
OTHER INFORMATION: EST_HUMAN HIT: A1350947.1, EVALU 1.00e-78
OTHER INFORMATION: SWISSPROT HIT: P30414, EVALU 1.00e-108
US-09-864-761-37733

Query Match 27.8%; Score 50; DB 10; Length 253;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 SGETPSEDDRSQS 29
Db 203 SDETPSRDDDSQS 217

Search completed: June 17, 2003, 12:02:30
Job time: 5.74085 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:06 ; Search time 4.04832 seconds
(without alignments)
831.137 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180

Sequence: 1 SFQSSSTPTSGFGSGKETPSEDNRQSREHMGES 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	33.9	974	2	A40580
2	56	31.1	621	2	T48187
3	55	30.6	89	2	B95010
4	55	30.6	89	2	H97881
5	54.5	30.3	2514	2	F81045
6	54	30.0	715	2	D96582
7	53.5	29.7	737	1	KIRTC
8	53.5	29.7	737	1	KIMSCE
9	53	29.4	384	2	T23604
10	53	29.4	390	2	G01936
11	53	29.4	814	2	T26702
12	53	29.4	3938	2	T42761
13	52	28.9	218	2	C86337
14	52	28.9	675	2	S51037
15	52	28.9	775	2	A61228
16	52	28.9	976	1	TVMSMD
17	52	28.9	978	1	macrophage colony-
18	52	28.9	2248	2	A35938
19	51.5	28.6	427	2	JN0785
20	51.5	28.6	2700	2	D88450
21	51	28.3	83	1	NINJ1F
22	51	28.3	92	2	AE3242
23	51	28.3	132	2	S22997
24	51	28.3	132	2	T08535
25	51	28.3	210	2	F72645
26	51	28.3	414	2	T34000
27	51	28.3	610	2	S41315
28	51	28.3	642	2	C86152
29	51	28.3	1711	1	A47392

30	51	28.3	1748	2	S42136	cnfB protein - Tet
31	50.5	28.1	430	2	T18651	hypothetical prote
32	50.5	28.1	522	2	S41819	nucleoporin p62 -
33	50.5	28.1	579	2	A84227	hypothetical prote
34	50	27.8	280	2	UC2358	melanoma antigen M
35	50	27.8	454	2	T29917	hypothetical prote
36	50	27.8	671	2	D72346	chemotaxis sensor
37	50	27.8	686	2	S70180	hypothetical serin
38	50	27.8	710	2	T41586	hypothetical prote
39	50	27.8	735	2	T45059	probable RNA helic
40	50	27.8	819	2	T08745	hypothetical prote
41	50	27.8	829	2	T29372	SERA antigen/papal
42	50	27.8	930	2	D71617	natural killer cel
43	50	27.8	1403	1	A47328	hypothetical prote
44	50	27.8	1770	2	A71517	vitellogenin I pre
45	50	27.8	1912	2	T29088	

ALIGNMENTS

RESULT 1
A40580
Iodestiar maternal-effect protein - fruit fly (Drosophila melanogaster)
N:Alternate names: probable nucleoside triphosphate binding protein Iodestiar
C:Species: Drosophila melanogaster
C>Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 21-Jul-2000
C:Accession: A40580; S19008
R:Girdham, C.H.; Glover, D.M.
Genes Dev. 5, 1786-1799, 1991
A>Title: Chromosome tangling and breakage at anaphase result from mutations in Iodest
A:Reference number: A40580; MUID:92009170; PMID:1916263
A:Accession: A40580
A:Molecule type: mRNA
A:Residues: 1-974 <GIR>
A:Cross-references: GB:X62629; NID:98187; PIDN:CA44496.1; PID:98188
C:Genetics:
A:Gene: Iodestiar
A:Cross-references: Flybase:FBgn002542

Query Match 33.9%; Score 61; DB 2; Length 974;
Best Local Similarity 51.9%; Pred. No. 7.1;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 4 SESSTPTSGFGSGKETPSEDNRQSRE 30
DB 39 SKSRPSSAGVVIDETQSEESQSSSE 65

RESULT 2

T48187
hypothetical protein F7A7.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48187
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.;

ews, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24487

A:Accession: T48187

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-621 <BEV>

A:Cross-references: EMBL:AL161946

A:Experimental source: Cultivar Columbia; BAC clone F7A7

A:Map position: 5
A:Introns: 65/3; 263/3; 289/3; 403/1; 469/3; 495/2; 529/2; 592/3
A>Note: F7A7.180

Query Match 31.1%; Score 56; DB 2; Length 621;
Best Local Similarity 38.7%; Pred. No. 20;
Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

OY 1 SF0ESSTPSTGCGKTPSEDDRSQSRHMG 31
 DB 475 SVEHKSSIVAVGQYDGKREYLTMAERFDPREH 505

RESULT 3

hypothetical protein SP0088 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: B95010
 R:Retellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: B95010
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-89 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK4275.1; PID:g14971554; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP0088

Query Match 30.6%; Score 55; DB 2; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.8;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 3 QSESSTPSTGCGKTPSEDDRSQSRHMG 34
 DB 55 ESOSATKDGQGTAGKETAGEDSANQTOEISQ 86

RESULT 4

hypothetical protein spr0080 [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: H97881
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: H97881
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-89 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK98884.1; PID:g15457615; GSPDB:GN00174
 C:Genetics:
 A:Gene: spr0080

Query Match 30.6%; Score 55; DB 2; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.8;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 3 QSESSTPSTGCGKTPSEDDRSQSRHMG 34
 DB 55 ESOSATKDGQGTAGKETAGEDSANQTOEISQ 86

RESULT 5

hemagglutinin/hemolysin-related protein NMB1768 [imported] - Neisseria meningitidis (str
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: F81045

R:Retellin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
 ri, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Masignani, V.; Piazza, M.
 Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.;
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: F81045
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2514 <TEF>
 A:Cross-references: GB:AE002526; GB:AE002098; NID:g7227015; PIDN:AAK42109.1; PID:g722
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1768

Query Match 30.3%; Score 54.5; DB 2; Length 2514;
 Best Local Similarity 36.4%; Pred. No. 1.3e+02;
 Matches 16; Conservative 9; Mismatches 8; Indels 11; Gaps 3;

OY 3 QSESST-----PSTG-GFS-GKETPSDDRSQSRHMGES 35
 DB 1537 QSQSETHYQOTKSGLSAGIGFTIGSKTNTQENQSNMHTGIST 1580

RESULT 6

hypothetical protein F1511.23 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: D96582
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maitl, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D96582
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-715 <STO>
 A:Cross-references: GB:AE005173; NID:g4587556; PIDN:AA025787.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F1511.23
 A:Map position: 1

Query Match 30.0%; Score 54; DB 2; Length 715;
 Best Local Similarity 36.8%; Pred. No. 42;
 Matches 14; Conservative 4; Mismatches 16; Indels 4; Gaps 1;

OY 1 SF0ESSTPSTGCGKTPSEDDRSQSRHMG 34
 DB 40 SASSTSTPSTGTLGLSLFSGASVSSSSSSSHPSVGE 77

RESULT 7

protein kinase C (EC 2.7.1.1) epsilon - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
 C:Accession: B28163; B26408; S00216
 R:Ono, Y.; Fujii, T.; Ogita, K.; Kikkawa, U.; Igarashi, K.; Nishizuka, Y.
 J. Biol. Chem. 263, 6927-6932, 1988
 A:Title: The structure, expression, and properties of additional members of the prote
 A:Reference number: A92717; MUID:88198270; PMID:2834397
 A:Accession: B28163
 A:Molecule type: DNA
 A:Residues: 1-737 <ONO>

A:Cross-references: GB:M1831; NID:g206182; PIDN:AAA41872.1; PID:g206183
 R:Housey, G.M.; O'Brien, C.A.; Johnson, M.D.; Kirschmeier, P.; Weinstein, I.B.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1065-1069, 1987
 A:Title: Isolation of cDNA clones encoding protein kinase C: evidence for a protein kinase
 A:Reference number: A64145; MUID:87147193; PMID:3469647
 A:Accession: B26408
 A:Molecule type: mRNA
 A:Residues: 397-447; 'GQGLHDDREDFGSGAET', 467, 'LSNPILLPLDPOGPELRLNQ', 487-545, 'C', 547-636
 A:Cross-references: GB:M15523; NID:g206192; PIDN:AAA41877.1; PID:g206193
 C:Comment: Protein kinase C epsilon and epsilon' appear to be encoded by the same gene
 C:Function:
 A:Note: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
 C:Superfamily: protein kinase C delta; protein kinase C zinc-binding repeat homology; pi
 C:Keywords: alternative splicing; ATP; autophosphorylation; duplication; phorbol ester
 F:156-161/Region: pseudophosphorylation motif
 F:170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
 F:243-292/Domain: protein kinase C zinc-binding repeat homology <K22>
 F:406-668/Domain: protein kinase homology <KIN>
 F:414-422/Region: protein kinase ATP-binding motif
 F:170-201-204-220/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
 F:189-186-209-212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
 F:243-273-276-292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
 F:256-259-281-284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
 F:437-456-532-534/Active site: Lys, Glu, Asp, Lys #status predicted
 F:703-710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match 29.7%; Score 53.5; DB 1; Length 737;
 Best Local Similarity 48.0%; Pred. No. 50;
 Matches 12: Conservative 5; Mismatches 3; Indels 5; Gaps 1;

OY 4 SSSSTPSTGFGSGKETPSEDDRSQS 28
 DB 327 AESPPASG-----NSPSEDDRSKS 346

RESULT 8
 KIMSC
 A:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 21-Nov-1997
 C:Accession: S02270
 R:Schnap, D.; Parker, P.J.; Bristol, A.; Kriz, R.; Knopf, J.
 FEBS Lett. 243, 351-357, 1989
 A:Title: Unique substrate specificity and regulatory properties of PKC-epsilon: a ration
 A:Reference number: S02270; MUID:89137541; PMID:2917656
 A:Accession: S02270
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-737 <SCH>
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
 A:Note: activity is calcium-independent, phospholipid-dependent, and activated by diacyl
 C:Superfamily: protein kinase C delta; protein kinase C zinc-binding repeat homology; pi
 C:Keywords: ATP; autophosphorylation; duplication; phorbol ester binding; phospholipid
 F:156-161/Region: pseudophosphorylation motif
 F:170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
 F:243-292/Domain: protein kinase C zinc-binding repeat homology <K22>
 F:406-668/Domain: protein kinase homology <KIN>
 F:414-422/Region: protein kinase ATP-binding motif
 F:170-201-204-220/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
 F:189-186-209-212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
 F:243-273-276-292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
 F:256-259-281-284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
 F:437-456-532-534/Active site: Lys, Glu, Asp, Lys #status predicted
 F:703-710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match 29.7%; Score 53.5; DB 1; Length 737;
 Best Local Similarity 48.0%; Pred. No. 50;
 Matches 12: Conservative 5; Mismatches 3; Indels 5; Gaps 1;

OY 4 SSSSTPSTGFGSGKETPSEDDRSQS 28

DB 327 AESPPASG-----NSPSEDDRSKS 346

RESULT 9
 T23604
 A:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T23604; T26115
 R:Kershaw, J.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: T23604
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-384 <WIL>
 A:Cross-references: EMBL:283115; PIDN:CAB05558.1; GSPDB:GN00019; CESP:K11D2.4
 A:Experimental source: clone K11D2
 R:Lenhard, N.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z20153
 A:Accession: T26115
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-384 <WIL>
 A:Cross-references: EMBL:281137; PIDN:CAB03472.1; GSPDB:GN00019; CESP:K11D2.4
 A:Experimental source: clone W02D9
 C:Genetics:
 A:Gene: CESP:K11D2.4
 A:Map position: 1
 A:introns: 46/1; 103/1; 159/3; 225/1; 296/1; 354/3

Query Match 29.4%; Score 53; DB 2; Length 384;
 Best Local Similarity 36.2%; Pred. No. 30;
 Matches 13: Conservative 3; Mismatches 18; Indels 0; Gaps 0;

OY 1 SF0ESSPTSTGFGSGKETPSEDDRSQREHMG 34
 DB 256 STSSAQPTHTGGGGGQTPSSASGSGSKD 289

RESULT 10
 G01936
 A:Title: binding protein 3 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G01936
 R:Ren, R.
 submitted to the EMBL Data Library, July 1995
 A:Reference number: G08875
 A:Accession: G01936
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-390 <REN>
 A:Cross-references: EMBL:U31089; NID:g987264; PID:g987265
 C:Genetics:
 A:Gene: Ab1BP3
 C:Superfamily: SH3 homology
 F:335-382/Domain: SH3 homology <SH3>

Query Match 29.4%; Score 53; DB 2; Length 390;
 Best Local Similarity 45.5%; Pred. No. 30;
 Matches 15: Conservative 2; Mismatches 14; Indels 2; Gaps 1;

OY 1 SF0ESSPTSTGFGSGKETPSEDDRSQREHMG 33
 DB 142 SVNORRTYSTSGSGSGSPS--SRSSRENSG 172

RESULT 11
 T26702
 hypothetical protein Y38H6C.14 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26702
 R:White, S.
 Submitted to the EMBL Data Library, September 1998
 A:Reference number: Z20255
 A:Accession: T26702
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-814 <MIL>
 A:Molecule type: DNA
 A:Cross-references: EMBL:AL031630; PIDN:CAA20992.1; GSPDB:GN00023; CESP:Y38H6C.14
 A:Experimental source: clone Y38H6C
 C:Genetics:
 A:Gene: CESP:Y38H6C.14
 A:Map position: 5
 A:Introns: 25/3; 370/2; 459/1; 505/3; 546/3; 666/2; 686/2

Query Match 29.4%; Score 53; DB 2; Length 814;
 Best Local Similarity 45.7%; Pred. No. 64;
 Matches 16; Conservative 4; Mismatches 7; Indels 8; Gaps 2;

QY 1 SFOSE-----SSTPTGSGFKETPSDDRSOS 28
 11:11 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 756 SFESEDKSEKSGSTPSESG-SGSESENEKSRSSNS 789

RESULT 12
 T42761
 Bassoon protein - rat
 N:Alternate names: brain-specific synapse-associated protein
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T42761
 R:Dieck, S.; Sammarti-Villa, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex, J.; Cell Biol. 142, 499-509, 1998
 A:Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized
 A:Reference number: Z22429; MUID:96345363; PMID:9679147
 A:Accession: T42761
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-3938 <DIE>
 A:Cross-references: EMBL:Y16563; NID:93413503; PIDN:CAA76287.1; PID:93413504
 A:Experimental source: strain Sprague Dawley; brain
 C:Function:
 A:Description: may be involved in cytomatrix organization at the site of neurotransmitter
 A:Note: component of the presynaptic cytoskeleton
 C:Keywords: coiled coil; zinc finger

Query Match 29.4%; Score 53; DB 2; Length 3938;
 Best Local Similarity 42.3%; Pred. No. 3.1e+02;
 Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 6 SSTPTGSGFKETPSDDRSQSRNH 31
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 3166 SSTPTGSGFKETPSDDRSQSRNH 3191

RESULT 13
 C86337
 hypothetical protein F14010.5 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C:Accession: C86337
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C86337
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE005172; NID:99558591; PIDN:AAF8154.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: mitochondrial inner membrane protein TIM17

Query Match 28.9%; Score 52; DB 2; Length 218;
 Best Local Similarity 37.8%; Pred. No. 23;
 Matches 14; Conservative 7; Mismatches 10; Indels 6; Gaps 1;

QY 5 ESSPTGSGFS-----GKETPSDDRSQSRNHGES 35
 11:11 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 170 ETSSSSGVSWMGSLFKKKKKEEDHSESTHILES 206

RESULT 14
 S51037
 zinc-finger protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Dec-2000
 C:Accession: S51037; S06565
 R:Schaefer, U.; Rausch, O.; Bouwmeester, T.; Pleier, T.
 Eur. J. Biochem. 226, 567-576, 1994
 A:Title: Sequence-specific recognition of a repetitive DNA element by a C(2)H(2) zinc
 A:Reference number: S51037; MUID:95094815; PMID:8001572
 A:Accession: S51037
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-675 <SCH>
 R:Nietfeld, W.; El-Barradi, T.; Mentzel, H.; Pleier, T.; Koester, M.; Poeting, A.; Kno J. Mol. Biol. 208, 639-659, 1998
 A:Title: Second-order repeats in Xenopus laevis finger proteins.
 A:Reference number: S05632; MUID:90040698; PMID:2509712
 A:Accession: S06565
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 85-448, K, 450-613 <NIE>
 A:Experimental source: clone X1CGF20-1
 C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
 C:Keywords: DNA binding; zinc finger

Query Match 28.9%; Score 52; DB 2; Length 675;
 Best Local Similarity 37.5%; Pred. No. 71;
 Matches 12; Conservative 8; Mismatches 10; Indels 2; Gaps 1;

QY 3 QSESSTPTGSGFKETPSDDRSQSRNHGE 34
 11 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 392 QSTHTSPSTE--FGVOTEDNHQSPKDHGE 421

RESULT 15
 A61228
 collagen alpha 2(IV) chain precursor - rabbit (fragments)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 17-Mar-1999
 C:Accession: A61228
 R:Yamaguchi, N.; Sato, N.; Ko, J.S.; Minomiya, Y.
 Invest. Ophthalmol. Vis. Sci. 32, 2924-2930, 1991
 A:Title: Cloning of alpha1(IV) and alpha2(IV) collagen cDNAs from rabbit corneal endo
 A:Reference number: A61228; MUID:92010685; PMID:1717338
 A:Accession: A61228
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-775 <YAM>
 C:Superfamily: collagen alpha 1(IV) chain
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

Query Match 28.9%; Score 52; DB 2; Length 775;
 Best Local Similarity 37.5%; Pred. No. 82;

Matches 15; Conservative 3; Mismatches 8; Indels 14; Gaps 2;

Oy 9 PS-TGSPSGKETP-----SEDDRSQREHME 34
|| ||||| | ||:|:|:|

Db 149 PSGTGTFPGSPGPGKPGKGPYALSEDDRDRTKGEIGE 188

Search completed: June 17, 2003, 12:01:12
Job time : 6.24832 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:41 ; Search time 31.0952 Seconds

(Without alignments)
690.935 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698

Sequence: 1 RIRAMAQGVFMIDTQCSPKT.....VROAVREWLGRVRYDIIV 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	4.7	2278	1	FAB1_YEAST
2	124	4.6	2150	1	SDC3_CAEEL
3	113	4.2	482	1	YSR2_CAEEL
4	110	4.1	829	1	CRCM_HUMAN
5	109.5	4.1	342	1	HXCA_HUMAN
6	108.5	4.0	872	1	S3B2_HUMAN
7	108.5	4.0	3703	1	ABR1_HUMAN
8	108	4.0	1912	1	VIT1_CHICK
9	107	4.0	1762	1	DPOQ_HUMAN
10	106	3.9	390	1	MEI1_XENLA
11	105	3.9	506	1	GTA_NPVAC
12	105	3.9	1214	1	BRF3_HUMAN
13	104.5	3.9	876	1	VPA1_CAEEL
14	103	3.8	1502	1	MRP6_RAT
15	103	3.8	3726	1	ABF1_MOUSE
16	102	3.8	457	1	CMGA_HUMAN
17	101.5	3.8	4486	1	DYH9_HUMAN
18	100.5	3.7	453	1	SSF2_YEAST
19	100	3.7	1268	1	VGIN_HUMAN
20	99.5	3.7	342	1	HXCA_MOUSE
21	99.5	3.7	449	1	ENGA_MYCPN
22	99.5	3.7	527	1	RAG2_MOUSE
23	99.5	3.7	649	1	TOP3_SALTI
24	99	3.7	1456	1	YNEO_YEAST
25	99	3.7	1785	1	BIG2_HUMAN
26	98.5	3.7	449	1	CMGA_BOVIN
27	98.5	3.7	960	1	LI36_CAEEL
28	98.5	3.7	986	1	DPOL_NPVBM
29	98	3.6	1171	1	Y637_HUMAN
30	97.5	3.6	743	1	MYBB_XENLA
31	97.5	3.6	2314	1	AKA6_RAT
32	97	3.6	477	1	MEI2_HUMAN
33	97	3.6	700	1	TRDN_CANFA

ALIGNMENTS

RESULT 1
FAB1_YEAST
ID FAB1_YEAST STANDARD: PRT: 2278 AA.
AC P34756:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable phosphatidylinositol-4-phosphate 5-kinase FAB1 (EC 2.7.1.68)
DE (1-phosphatidylinositol-4-phosphate kinase) (PI5K) (Ptdins(4)P-5-kinase) (diphosphoinositide kinase).
GN FAB1 OR YFR019W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95392039; PubMed=7663021;
RA Yamamoto A., Dewald D.B., Boronikov I.V., Anderson R.A., Emr S.D., Koshland D.;
RT "Novel PI(4)P 5-kinase homologue, Fab1p, essential for normal vacuole function and morphology in yeast.";
RL Mol. Biol. Cell 6:525-539(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / AB972;
RC MEDLINE=95400292; PubMed=7670463;
RA Sasaki Y., Naitou M., Hagihara H., Shibata T., Ozawa M., Murakami S.-I., Sasagawa M., Tsuchiya Y., Soeda E., Yokoyama K., Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SIMILARITY TO TCP1/CCT FAMILY.
RX MEDLINE=95314774; PubMed=7794526;
RA Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RT "Primary structure of the thermosome from Thermoplasma acidophilum.";
RL Biol. Chem. Hoppe-Seyler 376:119-126(1995).
CC - FUNCTION: CATALYSES THE PHOSPHORYLATION OF PHOSPHATIDYLINOSITOL-4-PHOSPHATE ON THE FIFTH HYDROXYL OF THE MYO-INOSITOL RING. TO FORM PHOSPHATIDYLINOSITOL-4,5-BIPHOSPHATE. REQUIRED FOR ENDOCYTIC VACUOLAR PATHWAY AND NUCLEAR MIGRATION. THE PRODUCT OF THE REACTION IT CATALYSES FUNCTIONS AS AN IMPORTANT REGULATOR OF VACUOLE HOMEOSTASIS PERHAPS BY CONTROLLING MEMBRANE FLUX TO AND/OR FROM THE VACUOLE.
CC - CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-monophosphate - ADP + 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate.
CC - SUBCELLULAR LOCATION: VACUOLE-ASSOCIATED (POTENTIAL).
CC - SIMILARITY: BELONGS TO THE PTDINS(4)P-5-KINASE FAMILY.
CC - SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U01017; AAA81360.1; -
DR EMBL: D50617; BAA09258.1; -
DR SGD: S0001915; FABI.
DR InterPro: IPR002498; PIP5K.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam: PF01363; FYVE; 1.
DR Pfam: PF01504; PIP5K; 1.
DR SMART: SM00064; FYVE; 1.
DR SMART: SM00350; PIP5K; 1.
DR PROSITE: PS0178; ZF_FYVE; 1.
DR TRANSFERASE; Kinase; Zinc-finger; Zinc.
FT ZN_FING 240 299 FYVE-TYPE.
FT DOMAIN 393 397 POLY-PRO.
FT DOMAIN 571 590 POLY-ASR.
FT DOMAIN 1808 1811 POLY-THR.
FT DOMAIN 1891 1897 POLY-GLN.
FT CONFLICT 2275 2275 R -> W (IN REF. 2).
SQ SEQUENCE 2278 AA; 257417 MW; 1A0A30E13165DE41 CRC64;

Query Match 4.7%; Score 126; DB 1; Length 2278;
Best Local Similarity 21.5%; Pred. No. 0.62; Mismatches 148; Indels 170; Gaps 27;

Matches 108; Conservative 77; Mismatches 148; Indels 170; Gaps 27;

QY 19 KTPNPF--DHAQSC--OLIEL-----PPDE--KPNQTKKSVSFREIYVS 58
DB 1330 KCPHFTKDHVKGFGYNDLVLRLEYSDELEVHELTPPRKIKMKPHIDIKLVE----- 1382
QY 59 LLSHQVL--LQNKDILLE--EPYKGPSPGEEKTIQVEALAGLRLISQNLAVIFDL 114
DB 1383 -LYTKLEIKINFGSVLSRLERIKLDSMTDKV-----LSGQAKTILEKSNATEBOK 1434
QY 115 L-----LDYRTAREFEDSPGKCLLKVKSGIGCANLYROSAMSENIYFHALVCVL-- 167
DB 1435 LMLQDLDFYADSPCDQHLPLNLYIK-----SLX-DKAVNNSTFALFAKSYLPSE 1484
QY 168 TNOETITAEQVKVLFEDDERSTSSQCSSEDEDFEETAQVSPPRGKERQWRAMP 227
DB 1485 TDISRTIRAKQKLFY-----DSSRKDSEDKKSLHDEKAKTRKPEKNE-----LPL 1530
QY 228 LSYQPVSNADWVWLYKRLHKLMELCNNTYIQMHLDLNCME-----EP-----PIRK 275
DB 1531 EGLKQVE-----KPKIDSKNTTENRDTNEPQNAVITTTTKD 1567
QY 276 DPFILPSFOSE--STPSTGGFGSKETPSEDDRSO--SREHMGESLSLKAAGDILLPSP 333
DB 1568 DTPILPISGTHLVTPSPASSVSSSLTPQTERPPISSGTSIMT----- 1613
QY 334 KVEKKDPSRKKEWMENAGNKITYMAADKT---ISKLMTEYKRRQ-----OHNLISA 381
DB 1614 ---HDSSTRP-----NIRKMSDSSLCGLASLANEYSKNNKVSCLATFEQMHDA 1661
QY 382 FPKFVKYKKEEPL-----GPRGDS--PLLRQPHLMQGGCRHSFSAPELLRDK 432
DB 1662 LSKFEELREKERERQLQKDKQAIRLQSTPIVELYKWKD-----AVDE 1706
QY 433 RPRSGTSSSL-SVSVRDAAQI 454
DB 1707 PLHSRSSGNNSSANVKTLEAPV 1729

RESULT 2

SDC3_CAEEL STANDARD; PRT: 2150 AA.
AC P34706:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein sdc-3.

GN SDC-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
OC Rhabdilitae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=93161411; PubMed=8431944;
RA Klein R.D., Meyer B.J.;
RT "Independent domains of the Sdc-3 protein control sex determination
and dosage compensation in C. elegans.";
RL Cell 72:349-364(1993).

CC -1- FUNCTION: CONTROLS BOTH SEX DETERMINATION AND X CHROMOSOME DOSAGE
CC COMPENSATION. THESE TWO FUNCTIONS ACT INDEPENDENTLY.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EMBRYONIC AND EARLY LARVAL
STAGES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M85149; AAA28144.1; -
DR PIR: S27802; S27802.

DR InterPro: IPR000822; Znf_C2H2.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.

DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; FALSE NEG.

KM Developmental protein; Zinc-finger; Metal-binding; DNA-binding;
KW Nuclear protein; Repeat.

FT DOMAIN 443 987 DOSAGE COMPENSATION DOMAIN 1.

FT DOMAIN 1508 1516 SEX DETERMINATION DOMAIN.

FT DOMAIN 2080 2105 DOSAGE COMPENSATION DOMAIN 2.

FT ZN_FING 2078 2105 C2H2-TYPE.

FT ZN_FING 2117 2141 C2H2-TYPE.

SQ SEQUENCE 2150 AA; 249954 MW; 7430D77AC784E46 CRC64;

Query Match 4.6%; Score 124; DB 1; Length 2150;
Best Local Similarity 20.6%; Pred. No. 0.8;

Matches 114; Conservative 63; Mismatches 166; Indels 210; Gaps 24;

QY 33 IIEPPEKPNQHTKKSVSFREIYVSL-----SHQVLLQ---NLVDILLEFVK 79
DB 1177 IERLSEDDLLDQGSSEDFEIEIGVLLGESVEYQVQSVTISRSTSFESLVEDPEE 1236
QY 80 GP-----SPGEEKTIQ-VPEAKIAGFLRYISQNLAVIFDLSDSYRTAREFEDTSPGK 133
DB 1237 HPQQLPVASERANNQIYVEVEGVSVPVTVNQ-----EENVISGE--P 1279
QY 134 LKKVSGIGCANLYROSAMSFNTYFHALVCVAVLTNQ-----ETITAEQ----- 177
DB 1280 TLQSSSIPSSSHIT-----TVDELLGTESPGEATETPAEESPKKSKSGT 1326
QY 178 -----YKVLFEEDERSTSSQCSSEDEDFEETAQVSPPRGKERQWRAMP--L 228
DB 1327 TRGRPKVKENLKKRIQPRROCKEAAHEPE-VVEEDQVPEVGEVHPVAPAPAO 1385
QY 229 SVQPVSNADWVWLYKRLHKLMELCNNTYIQMHLDLNCMEP-----PIFGDPFILPS 283
DB 1386 ETEPI-----EQIIEEDKVFETIIR-----LPL 1410
QY 284 FQSESTPSTGGFGSKETPSEDD-----RSQREHMG 315
DB 1411 FETSPVPAPEGNIPBSRAHSDDDQVYISSETDPNCPINLVQYQNDKLTAVQYSTEELG 1470
QY 316 ESLSLKAGG-----DLLPSPKVKKKPKSKKKEWMEN--AGNKITYMAA 359
DB 1471 EYGEIDEGAPSPSEIVVHDEVLDQVLPKPKSKKGRRRKKTPPHIARAKVET--- 1527

QY 360 DKTSKLTMT-----YKKRKOHNLSA-----FPKEVKEKGE-----LGPR 398
 DB 1528 --SISK--TEEIELAPPTQSRKRMANSEATPATRQKAKKEEENDSVSLVPE 1583
 QY 399 GDSGLRPPHLMDOGQMRHSFSGAPPELLRODKPRSGSGSLSVS-----446
 DB 1584 PEDLEHETPRGPHVGEGETPSLRTGRE-----STASSVKTSRKRLPLSKNNP 1632
 QY 447 VRDAEOIOAMTN 459
 DB 1633 VPRMRIOQOAGTN 1645

RESULT 3

YSR2_CAEEL
 ID YSR2_CAEEL STANDARD: PRT: 482 AA.
 AC 009950:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 54.6 kDa protein F59B10.2 in chromosome II.
 GN F59B10.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Lloyd C.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: 248716; GAB8601.1;
 DR WormRep: F59B10.2; CE01591.
 KW Hypothetical protein.
 FT DOMAIN 214 282 SER-RICH.
 FT DOMAIN 228 241 POLY-SER.
 FT DOMAIN 276 282 POLY-SER.
 FT DOMAIN 378 384 POLY-SER.
 FT DOMAIN 397 412 ARG-RICH.
 FT DOMAIN 404 412 POLY-ARG.
 SQ SEQUENCE 482 AA: 54641 MW: 1BD7D31E9DDAA74B CRC64:

Query Match 4.28: Score 113; DB 1; Length 482;
 Best Local Similarity 21.98; Pred No. 0.67; Indels 130; Gaps 16;

Matches 85: Conservative 46; Mismatches 128; Indels 130; Gaps 16;
 QY 118 SYTAREEDTS-----PGLKCLLKRVSGIGANLYROSAMFNIFYFHLVCAVLTN 169
 DB 81 SLKDIRELRLNSIDQSLIKPPELKMRLSKAAG-----ESKRHAHFESIQEEKIS 130
 QY 170 QETITAEQYKVLFEEDERSIDSSQCCSEDDIFEETAO---VPPPKKEKRORARAP 226
 DB 131 EDLPFAQLPSSSPF---KKAIOGSESSSDSIIFDEVFEVLPSPPR-----175
 QY 227 LLSVQVSNADVWVLVKKRLKLMELCNNTYQMHLLENCKMEPEPT-----FKGDPF 278
 DB 176 --KPAARTAPIV-----VEKKIEKPAVKEGKAKKKKEKPT 210
 QY 279 FLTPSFQ-----SESTPSTGFGSGKEPSEDDR---SOSREHGESISLAKGCD 326
 DB 211 PTSESSFESSSDSTSESSSSSSASESESKSESOVSSSKTSTSKASSSKAYGSD 270
 QY 327 LLLPSPVYEKKDPBRKKEMWENAGNKIYTMADTISKL---MTEYKKRKOQNLNLSAF 382

DB 271 F-----ESEKSSSS-----SASTISKVTPKKLKDQKTKPKDKKRAK 307
 QY 393 PKEVKEKGEPLGRGDSPLRQPHLMDOGQMRHSFSGAPELL-----428
 DB 308 PDDIRQNKPEPT-PEYDFIPRLSSRSSNSSESTVRETNRPLESLKTKLAINKSSEKME 366
 QY 429 --RODKR--PRSGTSGSLSVSVRDAEO 453
 DB 367 KPRKDIRRAPRSSSSSSS---TLRDAENE 392

RESULT 4

CRCM_HUMAN
 ID CRM_HUMAN STANDARD: PRT: 829 AA.
 AC P23508:
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Colorectal mutant cancer protein (MCC protein).
 GN MCC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=9116485; PubMed=1848370;
 RA Kinzler K.W., Nishisho I., Vogelstein B., Bryan T.M., Levy D.B.,
 RA Smith K.J., Preisinger A.C., Hamilton S.R., Hedge P., Markham A.,
 RA Carlson M., Joslyn G., Groden J., White R., Miki Y., Miyoshi Y.,
 RA Nishisho I., Nakamura Y.;
 RT "Identification of a gene located at chromosome 5q21 that is mutated
 RT in colorectal cancers."
 RL Science 251:1366-1370(1991).
 RN [2]
 RP VARIANTS COLORECTAL CANCER MUTANTS.
 RX MEDLINE=91335211; PubMed=1651563;
 RA Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,
 RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,
 RA Petersen G., Hamilton S.R., Nishisho I., Miki Y., Bryan T.M.,
 RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.;
 RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer
 RT patients."
 RL Science 253:665-669(1991).
 CC -1- FUNCTION: CANDIDATE FOR THE PUTATIVE COLORECTAL TUMOR SUPPRESSOR
 CC GENE LOCATED AT 5Q21.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES TYPES.
 CC -1- DISEASE: IS PROBABLY INVOLVED IN EARLY STAGES OF COLORECTAL
 CC NEOPLASIA IN BOTH SPORADIC AND FAMILIAL TUMORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M62397; AAA52069.1;
 DR PIR: A33166; A33166.
 DR PIR: A38434; A38434.
 DR Genew: HGNC:6935; MCC.
 DR MIM: 159350;
 KW Anti-oncogene; Disease mutation.
 FT VARIANT 267 267 R -> L (IN COLORECTAL CANCER).
 FT VARIANT 486 486 P -> L (IN COLORECTAL CANCER).
 FT VARIANT 490 490 S -> L (IN COLORECTAL CANCER).
 FT VARIANT 506 506 R -> Q (IN COLORECTAL CANCER).
 FT FTID-VAR_005144.

FT VARIANT 698 698 A -> V (IN COLORECTAL CANCER).
 FT SEQUENCE 829 AA; 93055 MM; 5D45225FD91CA18F CRC64;
 Query Match 4.1%; Score 110; DB 1; Length 829;
 Best Local Similarity 18.4%; Pred. No. 2.3; Indels 130; Gaps 19;
 Matches 92; Conservative 81; Mismatches 198;
 OY 81 PSPEERTI-----OVPEAKLAGFLRYISMOMLAIVFDLLDSYRTAREFTSPGKILK 136
 DB PSTGELSTSSSSNDIPAKIAERVK-----LKTRESSSSDRPVLS 349
 OY 137 KVSIGGCAANLYROSASF-----NIFYHALVCANLTNOETITTAQVKVLFED 185
 DB EISSIGSVSAEHLASHLSDSCNIOELFOTLYSHG-----SAISEKIRE--FEV 398
 OY 186 DERSTDSQCCSDEDEFEETAAQVPPRGKRRWRAMPPLSVQPSNADWLVRL 245
 DB 399 ETEPLNRRIEHLKSONDLITLLE-----ECKSMERNMVLGKYSNLTALRLALQY 451
 OY 246 HKLCMELCNNTYOMHLLENCEMEPPFPKDPFLLPSQSESTPTSGFGSKETPSE- 304
 DB 452 SEQCIEAVYELL-----ALAESEQSLILGCFRAAGVGSPPGDSGDENTOM 498
 OY 305 ----DSSQSRHMGESLSLKAGG--GDLLLPSPKVEKRPSPKKEWENAKKITVMA 358
 DB 499 LKRAHDCRKTAKENAKALMLDSCGAPFAVACSVOP-----WESLSSNSHTST 549
 OY 359 ADKTISKLMTEYKRRKQOHNLSAPFKVEYKKG-----EPLG----PRGOD 401
 DB 550 TSSSTASSODTEFTKEDEE-RLKDYIQOLKNDRAVKLTMLESHIDPLSYDVKPRDS 608
 OY 402 SPL-LQRP---QHLMDOQMHSFSAGPELLRQDKRPRSGSTGSSLSVDAEAQIQAW 457
 DB 609 QRLDENAVLMOELMAKEMAEKLAQYLLEKEK-----ALEIKLSTREAQEQAY 660
 OY 458 TNMVLTYINOIQLPDQFTALQ-----PAFPCIS----QLTCHVDIR 498
 DB 661 LVHIEHLSEVEEKEORMRSLSTSSGSKDKPEKECADASPALISLAELFTTSENELA 720
 OY 499 VR--QAVREWLGRVRYDII 517
 DB 721 AEFNNAIRREKKLKAQVEIV 741
 RESULT 5
 HXCA_HUMAN STANDARD; PRT; 342 AA.
 AC 09NYVD6; 015220; 09BVD5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-C10.
 GN HOXC10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 OX 111
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20296799; PubMed=10835276;
 RA de Stanchina E., Gabbellini D., Norio P., Giacca M., Peverall F.A.,
 RA Riva S., Falaschi A., Biamonti G.;
 RT "Selection of homeotic proteins for binding to a human DNA replication
 RT origin";
 RT J. Mol. Biol. 299:667-680(2000).
 RN 121
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN 131
 RP SEQUENCE OF 94-106 AND 258-297 FROM N.A.

RX MEDLINE=98019011; PubMed=9357979;
 RA Flagiello D., Gibaud A., Dutrillaux B., Poupon M.F., Malfoy B.;
 RT "Distinct patterns of all-trans retinoic acid dependent expression of
 RT HOXB and HOXC homeogenes in human embryonal and small-cell lung
 RT Carcinoma cell lines";
 RL FEBS Lett. 415:263-267(1997).
 CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEBOX FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC
 DR EMBL; AF255675; AAF67759.1; -
 DR EMBL; BC001293; AA01293.1; -
 DR EMBL; X99684; CA67999.1; -
 DR EMBL; X99685; CA68000.1; -
 DR HSSP; P02833; 9AAT.
 DR TRANSFAC; T03328; -
 DR Genew; HGNC:5122; HOXC10.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox.1.
 DR ProDom; PD000010; Homeobox.1.
 DR SMART; SM00389; HOX.1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT DNA_BIND 268 327 HOMEBOX.
 FT CONFLICT 118 118 K -> N (IN REF. 1).
 FT CONFLICT 265 265 A -> G (IN REF. 3).
 FT CONFLICT 271 271 MISSING (IN REF. 3).
 FT SEQUENCE 342 AA; 38072 MM; BD8127FD43CA2A37B CRC64;
 Query Match 4.1%; Score 109.5; DB 1; Length 342;
 Best Local Similarity 19.3%; Pred. No. 0.76; Indels 121; Gaps 11;
 Matches 64; Conservative 38; Mismatches 108;
 OY 144 AANLYROSAMSFNIFYHALVCANLTNOETITAEQVKVLFEDDERSTDSQCCSDEDI 203
 DB 28 SAGWTMGSGSDFN-----CGVM-----RGGLAPSLSKRDEGS 60
 OY 204 FEETAAQVSPR-----GKEKQWRAMPPLSVQPSNADWLVRLKHLKLCMELCNN 255
 DB 61 SPSTALNTYPSYLSQLDSWGDPAKAYRLEQPV--GRPLSSCSPYPPEEENVCNYSAEK 118
 OY 256 ----YIOMHDLNCEMEPPFPKDPFLLPSQSESTPTSGFGSKETPSPDDR 307
 DB 119 RAKGPEPALYSHLPESCIEGHEVPVSYRASPSYALDKTCHSGANFEAPFQRA 178
 OY 308 SOS--REHMGESLSLKAGGDLPLPSPKVEKRPSPKKEWENAGNKITYMAADKITYSK 365
 DB 179 SLNRAEHL-ESPQL--GKAVSFPEPKKSQSPS----- 210
 OY 366 LMTEYKRRKQOHNLSAPFKVEYKKGPELLRGQDSPLQRPQHLMDQGMHSFSAGP 425
 DB 211 -----PNEIKTEQ--SLADPKGSPS----- 228
 OY 426 ELLRQDKRPRSGSTGSSLSVDAEAQIOA 456
 DB 229 ----ESEKERAKADSSPTSDNEAKEIKRA 255
 RESULT 6
 S3B2_HUMAN

```

ID   S3B2_HUMAN          STANDARD:          PRT:      872 AA.
DC   013435.
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Splicing factor 3B subunit 2 (Spliceosome associated protein 145) (SAP
GN   S3B2 OR SAP145.
OS   Homo sapiens (human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX   NCBI_TaxID:9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 151-159 AND 794-817.
RX   MEDLINE=96154048; PubMed=8566756;
RA   Gorzani O., Feld R., Reed R.;
RT   "Evidence that sequence-independent binding of highly conserved U2
RT   snRNP proteins upstream of the branch site is required for assembly
RT   of spliceosomal complex A."
RL   Mol. Cell Dev. 10:233-243(1996).
RN   [2]
RP   CHARACTERIZATION OF THE SPLICOSOME.
RX   MEDLINE=20337962; PubMed=10882114;
RA   Das R., Zhou Z., Reed R.;
RT   "Functional association of U2 snRNP with the ATP-independent
RT   spliceosomal complex E."
RL   Mol. Cell 5:779-787(2000).
CC   -1- FUNCTION: SUBUNIT OF THE SPLICING FACTOR SFB3 REQUIRED FOR 'A'
CC   BRANCHPOINT SEQUENCE (BPS) IN PRE-MRNA. SEQUENCE INDEPENDENT
CC   BINDING OF SFB3/SFB3 COMPLEX UPSTREAM OF THE BRANCH SITE IS
CC   ESSENTIAL, IT MAY ANCHOR U2 SNRP TO THE PRE-MRNA. MAY ALSO BE
CC   INVOLVED IN THE ASSEMBLY OF THE 'E' COMPLEX. BELONGS ALSO TO THE
CC   MINOR U12-DEPENDENT SPLICOSOME, WHICH IS INVOLVED IN THE SPLICING
CC   OF RARE CLASS OF NUCLEAR PRE-MRNA INTRON.
CC   -1- SUBUNIT: COMPONENT OF SPLICING FACTOR SFB3 WHICH IS COMPOSED OF
CC   FOUR SUBUNITS, SFB34/SNP49, SFB33/SNP130, SFB32/SAP145,
CC   SFB31/SAP135. SFB3 ASSOCIATES WITH THE SPLICING FACTOR SFB3A AND A
CC   12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS
CC   COMPLEX (U2 SNRNP). SFB32 INTERACTS DIRECTLY WITH SFB34.
CC   -1- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC   -1- SIMILARITY: TO YEAST CUS1. SOME TO C.ELEGANS ZK632.11.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on ways
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL; U41371; AAA97461.1; .
DR   Genew; HGNC:10769; SFB32.
DR   MIM; 605591; .
DR   InterPro; IPR003034; SAP.
DR   Pfam; PF02037; SAP: 2.
DR   SMART; SM00513; SAP: 1.
KW   Spliceosome; mRNA processing; mRNA splicing; Nuclear protein.
FT   DOMAIN            68      73      POLY-PRO.
FT   DOMAIN            81      90      POLY-PRO.
FT   DOMAIN            106     109     POLY-PRO.
FT   DOMAIN            226     230     POLY-PRO.
FT   DOMAIN            269     274     POLY-GLU.
FT   DOMAIN            308     312     POLY-LYS.
FT   DOMAIN            676     679     POLY-GLU.
FT   DOMAIN            697     703     POLY-GLU.
SQ   SEQUENCE 872 AA; 97656 MW; AED669FDDA5DE31 CRC64;
Query Match 4.0%; Score 108.5; DB 1; Length 872;
Best Local Similarity 19.4%; Pred. No 3.2; Mismatches 97; Gaps 13;
Matches 63; Conservative 46; Indels 97; Gaps 13;

```

Db	259	SROEENMSQ-----EEEMETDTRSSLGOSASETEEDIVSYSKKKRNRNRNRKKKK	312
Qy	220	OMFARNPPLLISVOPVSNADVMVLVKRLHLCMELCNNYIOMHLDJENCMEEPPIFKGDPP	279
Db	313	PQVVR---GVSESSGCD-----REKSDTRSGSDSPADVAEIVEYTEPEEIEPPIF	362
Qy	280	ILPSFS-----ESSTGTGFSCKETPESDDDSQSEMHGES	317
Db	363	FKRIEAFKLTLDVKKKEKEKEPEKLDKLENSAAPKKGFEEHNDSDSDSDEQK---	419
Qy	318	LSLAKAGGDLPLPSPKYVEKDKPSRKKEMMENAGNKITYMAADTKISKLMEYKKRKROH	377
Db	420	-----KPEAPKSKKRLRMNR-----FTVAE-----LKQLVARDVDEM	455
Qy	378	NLSA-FPK---EYKVEKGEPL-----GPRGDSPLQRPQHLMDOG--OMRH	419
Db	456	DVTAQDPKLTVLHLKATRNVPVPRHMCERKRYLOGKRGIEKPPPELPDIKRTGIOEMRE	515
Qy	420	SFSAGE-----LLRODKRPRSG	437
Db	516	ALQKEEQKTKMSKREKRYPRMG	539

RESULT 7

ABF1_HUMAN	STANDARD:	PRT: 3703 AA.
ID	ABF1_HUMAN	015911: 013719; 015101;
AC	16-OCT-2001 (Rel. 40, created)	
DT	16-OCT-2001 (Rel. 40, last sequence update)	
DT	16-OCT-2001 (Rel. 40, last annotation update)	
DE	Alpha-fetoprotein enhancer binding protein (AT motif-binding factor) (AT-binding transcription factor 1).	
GN	ATF1.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	SEQUENCE FROM N.A. (ISOFORM A).	
RP	TISSUE=Lung;	
RC	MEDLINE=96070776; PubMed=7592926;	
RX	Miyata Y., Tam T., Ido A., Morinaga T., Miki T., Hashimoto T.,	
RA	Tamaoki T.;	
RT	"Cloning and characterization of an ATF1 isoform that expresses in a	
RL	neuronal differentiation-dependent manner."	
RL	J. Biol. Chem. 270:26840-26848(1995).	
RN	[2]	
RP	SEQUENCE FROM N.A. (ISOFORM B).	
RC	TISSUE=Hepatoma;	
RC	MEDLINE=92049333; PubMed=1719379;	
RA	Morinaga T., Yasuda H., Higashio K., Tamaoki T.;	
RT	"A human alpha-fetoprotein enhancer-binding protein, ATF1, contains	
RL	four homeodomains and seventeen zinc fingers."	
RL	Mol. Cell. Biol. 11:6041-6049(1991).	
RN	[3]	
RP	SEQUENCE OF 1-1190 FROM N.A. (ISOFORM A).	
RC	MEDLINE=99425270; PubMed=10493829;	
RA	Loftus B.J., Kim U.-J., Sheddov V.P., Kalush F., Brandon R.,	
RA	Psilman J., Mason T., Crosby M.L., Barnstead M., Cronin L.,	
RA	DeSilvas Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,	
RA	Etchler E.E., Harris P.C., Venter J.C., Adams M.D.;	
RT	"genome duplications and other features in 12 Mb of DNA sequence from	
RL	human chromosome 16p and 16q."	
RL	Genomics 60:295-308(1999).	
RN	[4]	
RP	SEQUENCE OF 1151-3703 FROM N.A.	
RA	Kozlowski A., McQuerry Y., Hotte M.;	
RL	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.	
CC	-!- FUNCTION: Transcriptional activator that binds to the AT-rich core	
CC	sequence of the enhancer element of the Afp gene.	
CC	-!- SUBCELLULAR LOCATION: Nuclear.	
CC	-!- ALTERNATIVE PRODUCTS: 2 isoforms: A (shown here) and B; are	

CC	Produced by alternative splicing.	
CC	-1- SIMILARITY: CONTAINS 4 HOMEBOX DOMAINS	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; L32832; AAC14462.1; -	
DR	EMBL; D10250; BAA01095.1; -	
DR	EMBL; AC002044; AAC31674.1; -	
DR	EMBL; AC004943; AAC79153.1; -	
DR	HSSP; P20263; 10CP	
DR	TRANSFAC; T00048; -	
DR	TRANSFAC; T01665; -	
DR	GeneW; HGNC:777; ATBFL	
DR	MIM; 104155; -	
DR	InterPro; IPR001356; Homeobox	
DR	InterPro; IPR000822; znf.C2H2	
DR	InterPro; IPR003604; znf.U1	
DR	Pfam; PF000046; homeobox; 4	
DR	Pfam; PF00096; zf-C2H2; 20	
DR	ProDom; PD000010; Homeobox; 4	
DR	SMART; SM00389; HOX; 4	
DR	SMART; SM00355; znf.C2H2; 21	
DR	SMART; SM00451; znf.U1; 7	
DR	PROSITE; PS00027; HOMEBOX_1; 2	
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14	
DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 9	
KW	Transcription regulation; Activator; Zinc-finger; Metal-binding;	
KW	DNA-binding; Homeobox; Nuclear protein; Repeat; Polymorphism;	
KW	Alternative splicing	
FT	ZN_FING	282
FT	ZN_FING	305
FT	ZN_FING	640
FT	ZN_FING	671
FT	ZN_FING	694
FT	ZN_FING	726
FT	ZN_FING	804
FT	ZN_FING	828
FT	ZN_FING	945
FT	ZN_FING	968
FT	ZN_FING	1008
FT	ZN_FING	1040
FT	ZN_FING	1064
FT	ZN_FING	1088
FT	ZN_FING	1112
FT	ZN_FING	1223
FT	ZN_FING	1246
FT	ZN_FING	1252
FT	ZN_FING	1275
FT	ZN_FING	1360
FT	ZN_FING	1385
FT	ZN_FING	1401
FT	ZN_FING	1423
FT	ZN_FING	1429
FT	ZN_FING	1452
FT	ZN_FING	1545
FT	ZN_FING	1569
FT	ZN_FING	1596
FT	ZN_FING	1620
FT	ZN_FING	1983
FT	ZN_FING	2006
FT	DNA_BIND	2145
FT	DNA_BIND	2204
FT	DNA_BIND	2242
FT	ZN_FING	2301
FT	ZN_FING	2328
FT	ZN_FING	2351
FT	ZN_FING	2530
FT	ZN_FING	2552
FT	DNA_BIND	2641
FT	ZN_FING	2700
FT	ZN_FING	2711
FT	DNA_BIND	2734
FT	ZN_FING	2944
FT	ZN_FING	3003
FT	ZN_FING	3024
FT	ZN_FING	3048
FT	ZN_FING	3529
FT	ZN_FING	3553
FT	ZN_FING	104
FT	ZN_FING	107
FT	DOMAIN	460
FT	DOMAIN	489
FT	DOMAIN	770
FT	DOMAIN	784
FT	DOMAIN	1723
FT	DOMAIN	1743
FT	DOMAIN	1789
FT	DOMAIN	1794
FT	DOMAIN	1852
FT	DOMAIN	1857
FT	DOMAIN	2037
FT	DOMAIN	2052
FT	DOMAIN	3197
FT	DOMAIN	3209
FT	DOMAIN	3210
FT	DOMAIN	3214
FT	DOMAIN	3227
FT	DOMAIN	3231
FT	DOMAIN	3376
FT	DOMAIN	3389
FT	DOMAIN	3396
FT	DOMAIN	3400
FT	DOMAIN	3404
FT	DOMAIN	3408
FT	DOMAIN	3412
FT	DOMAIN	3416
FT	DOMAIN	3420
FT	DOMAIN	3424
FT	DOMAIN	3428
FT	DOMAIN	3432
FT	DOMAIN	3436
FT	DOMAIN	3440
FT	DOMAIN	3444
FT	DOMAIN	3448
FT	DOMAIN	3452
FT	DOMAIN	3456
FT	DOMAIN	3460
FT	DOMAIN	3464
FT	DOMAIN	3468
FT	DOMAIN	3472
FT	DOMAIN	3476
FT	DOMAIN	3480
FT	DOMAIN	3484
FT	DOMAIN	3488
FT	DOMAIN	3492
FT	DOMAIN	3496
FT	DOMAIN	3500
FT	DOMAIN	3504
FT	DOMAIN	3508
FT	DOMAIN	3512
FT		

FT	DNAIN	3392	3395	POLY-GLN.
FT	DNAIN	3507	3527	POLY-GLX.
FT	DNAIN	3597	3600	POLY-PRO.
FT	DNAIN	3636	3639	POLY-SER.
FT	VARSPLIC	1	914	MISSING (IN ISOFORM B).
FT	VARIANT	3374	3374	A -> V.
FT			/FtId-VAR_011694.	
FT	VARIANT	3377	3384	MISSING.
FT			/FtId-VAR_011695.	
FT	VARIANT	3527	3527	G -> GGG.
FT			/FtId-VAR_011696.	
FT	CONFLICT	72	72	A -> S (IN REF. 3).
FT	CONFLICT	422	422	A -> P (IN REF. 3).
FT	CONFLICT	579	579	T -> A (IN REF. 3).
FT	CONFLICT	767	767	I -> S (IN REF. 3).
FT	CONFLICT	777	777	A -> V (IN REF. 3).
FT	CONFLICT	846	849	HHRV -> RHUG (IN REF. 3).
FT	CONFLICT	997	997	A -> S (IN REF. 3).
FT	CONFLICT	1150	1190	EALVDGSPETADPEAEKAKDQEGASSQAOKELTDSR -> GEWSRHMRPRLGCVHLLETSRCGLTFEGDVTPAGPH VPY (IN REF. 3).
SEQ	SEQUENCE	3703 AA:	404468 MW;	0f62AF37dAdCf856 CRC64;
<hr/>				
Query Match	Best Local Similarity	4.0%;	Score 108.5;	DB 1; Length 3703;
Matches	73; Conservative	45;	Mismatches	91; Indels 145; Gaps 18;
Oy	132	KCLKKVSGIGGANLYRQSAMSFNIEFHAVLCVLT-----	NOETT	174
Db	1072	KHLDQHSEVGDE-----SCYYHCYLCTSKAKLNLIGHVSKHQRSLSLK		1120
Oy	175	AEQVKVLFEDDE-----RSTD-----	SSQ	194
Db	1121	LQRLOKGIPEDEDLGQIFTRCPSPDPBEAIEDVEGPSETADPEELAKDQEGASSS		1180
Oy	195	OCSSDEDDIPEETNQVSPPRKKEKQRRAPPLISVOP-----	VSNA	236
Db	1181	QAEKELTDSPATSKRISPSSSES-----PLSSKRPKTAEDIKTPEOMYTQCPCKYSNA		1233
Oy	237	DWVWLVRKLH-----KCLMELCNN--YTONMLD-----LENCMDEEPIFK		274
Db	1234	D----VARLRHAMTQHSYOPMLRCPCODMLNNKHLDLHTLTHLSVAPDDCYEKLMTV		1289
Oy	275	GDFPFILPSFO-SSSTPTSTGGFGSKETPSBDRDSGRREIMGESLSLAKG--GGDLLP		330
Db	1290	TTPPMVMSSEFLPAAVPDNRGNNSMLEAG--OPEITSELGKNILPSASTEOSGD--LK		1345
Oy	331	PSPEVERKKSDSRKKE-----WMENAGKIKITMAADKITSLKMTYEK---KRRQ		375
Db	1346	PSF---ADPGSVREDSGFTCWKKKGQNVF-----KTSALALQHFNFVNHA KRQ		1390
<hr/>				
RESULT 8				
VITL_CHICK	STANDARD;	PRT;	1912 AA.	
AC	P87498;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUN-1998 (Rel. 36, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Vitellogenin I precursor (Minor vitellogenin) [Contains: Lipovitelin I (LVII); Phosvitin (PV); Lipovitelin II (LVIII); YGP42].			
GN	VTE1 OR VTGI.			
OS	Gallus gallus (Chicken).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Mabuchi N., Yamamura J.-I., Adachi T., Aoki N., Nakamura R.,			
RA	Matsuda T.;			
RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.			

FT CONFLICT 1756 1762 ELKDPDV -> RAKGL (IN REF. 2).
 SQ SEQUENCE 1762 AA: 197596 MM: AECDD1A2103F6BDA CRC64;
 Query Match 4.0%; Score 107; DB 1; Length 1762;
 Best Local Similarity 21.2%; Pred. No. 10;
 Matches 101; Conservative 71; Mismatches 163; Indels 142; Gaps 26;

OY 16 CSPKTPNNFHAQSCQLI-----TELPPDEKPNHTRKKSVSFRELIVSLSHOYLLONLY 70
 DB 722 CSNDESIIFSEMDGVQWALNDVDFVQDEKN-HT-----VVSPPRAL--ELS 766
 OY 71 DILLEEFKGPSPEEKTIQVPEAKLGFGLYISMONTAVIFDILLDSRYRAREEDPSG 130
 DB 767 DPVLDHHQDQDGGDDDERAKSKLTG-----TRQNSHFW-----SGASDLSFG 813
 OY 131 LKCLLKRVSGIGGANLYRQGSAMFNITYHAL-----VCAVLINQETITAEQYKV 181
 DB 814 LQRLIDKVS-----SPLENEKLSMTINFSLNKRKNTPELNEQEVISLLET---KQVGI 865
 OY 182 LF-----EDDERSTDSQCCSDEDIEFETAOVSPPRCKEKRQMRARAPL 228
 DB 866 SFSSNNVEKSKIEMLENNANHDETSLPRKESNIVDNGLI-PPTPIPTASAKLTPEGI 924
 OY 229 SVQPSVADYVWLVRKRLKLMELCNNYIQHLDLENCEBPPIFKGDPFELPSFOSES 288
 DB 925 LETEVN--PW-----KTNVYLQ-----PGSYLF-GSPSDIK---NHD 957
 OY 289 STPTGFGSGKETPSEDDRSQSRHMGESLSLAKGGDLLPSPKYEKK---DPSRK- 343
 DB 958 SPGRNGFRK-DNSPISD-----TSFSLQSQDQLQTLTPASSSESLSIIDVADQ 1006
 OY 344 -----KEMWENANKITYTMAADK-----TISKLMTEKKRKKQCHNLAFKYEKVEKKG 392
 DB 1007 NLFQFTIKEMCKKRFSL-SLACEKINSLTSSKTATIGSRKQ-----ASSQELPIRDQ 1061
 OY 393 EPLPGGDSPL-----ORPHLMDQGMHSP-----SAGPELLRQDK 432
 DB 1062 FPI--KGGDDTLVGLAVCMGGRDAYFSLQEKQKHSISASIVPSLDPLTLTKDR 1116

RESULT 10
 MEIL_XENLA STANDARD: PRT: 390 AA.
 AC P79937; P79938;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeobox protein Meis1 (Xmeis1).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 CC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97202105; PubMed=9049632;
 RA Steelman S., Moskow J.J., Muznyk K., North C., Druck T.,
 RA Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
 RT "Identification of a conserved family of Meis1-related homeobox
 genes";
 RL Genome Res. 7:142-156(1997).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: MEIS1-1 (SHOWN HERE)
 CC AND MEIS1-2: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE TALE/MEIS HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U68386; AAB19196.1;
 DR EMBL: U68387; AAB19197.1;
 DR HSSP: P41778; 1D06.
 DR TRANSFAC: T03390;
 DR TRANSFAC: T03391;
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; Homeobox; 1.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX_1; FALSE_NEG.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 DR DNA-binding; Nuclear protein; Homeobox; Alternative splicing.
 FT DOMAIN 194 240
 FT DOMAIN 242 269
 FT DOMAIN 262 267
 FT DNA_BIND 272 334
 FT VANSPLIC 373 390

FT
 FT
 FT VARIANT 34 34 K -> Q (IN MEIS1-2).
 FT VARIANT 91 91 F -> L (IN MEIS1-2).
 FT VARIANT 178 178 E -> K (IN MEIS1-2).
 FT VARIANT 273 273 N -> H (IN MEIS1-2).
 FT VARIANT 275 275 G -> K (IN MEIS1-2).
 FT VARIANT 315 315 MISSING (IN MEIS1-2).
 SQ SEQUENCE 390 AA: 42887 MM: DDD98637D08CF32P CRC64;
 Query Match 3.9%; Score 106; DB 1; Length 390;
 Best Local Similarity 19.2%; Pred. No. 1.6;
 Matches 56; Conservative 51; Mismatches 94; Indels 90; Gaps 12;

OY 181 VLFEDEKSTDSQO-----CSSE--DEDIETEAQVSPPRCKEKRQMRARAPL 230
 DB 89 LIFKCELATCTPREPGVAGDVCSSESFNEDIAVFSKQI-----RAEKPLSS 137
 OY 231 QPVSADPWVLY-----KRLHKCMELCNNTI-----QNHLDL-----ENKME 268
 DB 138 NPELDNLMITQAIQVLRFLLELEVHELCDFNCHRYTISCLGKMPIDLDVDDRQGSKD 197
 OY 269 EPIIFKDPFFILPSFOSE-----SSTPTGSGSGKETPSEDDRSQSRHMGES 317
 DB 198 SEDLTRSAPRLDQGSWSDHDDAASIRSGTIPRSGSGHT---SHSGNSSEQDGDG 254
 OY 318 LSLKAGGDDLPPSPKYEKKDPSRKKEWENAGKITYTMAADTISKLMTEYKKRQCH 377
 DB 255 IASPSTGDD-----DDPPKEKK--RNKGKIGFKVATNIMRAWLF-----QH 294
 OY 378 NLSAFPKEKVEKKEGPELG-----PRGDSPLLRQPHLMDQ 415
 DB 295 LTHPIPEQKKQLAQLDGLITLTVNNMFINARRIYQPMIDQSNRAVSOG 345

RESULT 11
 GYA_NPVAC STANDARD: PRT: 506 AA.
 AC P41447;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Probable global transactivator.
 GN GYA.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 CC NCB1_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RX Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear

RT polyhedrosis virus.
 RL Virology 202:586-605(1994).
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L22858; AAA66672.1;
 DR Interpro: IPR001410; DEAD.
 DR Interpro: IPR001650; Helicase_C.
 DR Interpro: IPR000330; SNF2_N.
 DR Pfam: PF00176; SNF2_N.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICC; 1.
 KM Helicase: ATP-binding.
 FT NP_BIND 58 64
 FT SITE 157 160 DEAD BOX.
 FT SEQUENCE 506 AA; 59058 MW; 84EDCC2A85984AE CRC64;

Query Match
 Best Local Similarity 22.3%; Score 105; DB 1; Length 506;
 Matches 69; Conservative 49; Mismatches 91; Indels 100; Gaps 18;

QY 126 DTSPGKCLIKVSGIGGAANLYR-QSAMSFNIFHALVCALVLTNOETTAEOVKVLF 184
 DB 222 DSTNRKISITIKIV-----LKRDKSEISSNIPKH-----TVEYV-HVNFN 260
 QY 185 DDERSTSSQCGSEDEDEETPAQVSPRGKEKRRARAPLLSVQPVSNADWVLYKR 244
 DB 261 EEEETLYDKLKCESE-----AYKAAVAARENALSR--LQOMOHV-----LWILK 306
 QY 245 LHKRMELCNVYIOMHLDLNCMEEPPIFKGDPF-----FILPS 283
 DB 307 LRQIC---CHPYLAMH--GKNILETNCDFKMDVSSCKRYLDVLDILMTSNDKILVS 361
 QY 284 -----FOS--ESSPTSGFGSGK-----EPSPEDDSOSREHGESLSLRAG 324
 DB 362 QWVEYLKIFENFQKNIAITLMTYGLKVEDRIIAETFTND--AANTQHRILLISIRCG 419
 QY 325 GDV-----LLPP--SPKVEKKDPSRKKEWENGNKIYTN--AADTISKLMEYK 371
 DB 420 VGLNLIGGNHIVMLPEHWPQIELQADDRISRMGQTKNTYVKMLNVEDNSIEK-----YI 475
 QY 372 KRKQOHNS 380
 DB 476 KQRQDKKIA 484

RESULT 12
 BRP3_HUMAN
 ID BRP3_HUMAN STANDARD; PRT; 1214 AA.
 AC 090LD4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 40, Last annotation update)
 DE Bromodomain and PHD finger-containing protein 3 (Fragment).
 GN BRP3 OR KIA1286.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleia; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20039619; PubMed=10574462;
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
 Ohara O.,

RT "Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res 6:337-345(1999).
 CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS 1 PWWP DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB033112; BA086600.1;
 DR HSSP: Q92831; B91.
 DR GeneW: HGNC:14256; BRP3.
 DR Interpro: IPR001487; Bromodomain.
 DR Interpro: IPR000313; PWWP domain.
 DR Interpro: IPR001965; Znf_PHD.
 DR Pfam: PF00439; Bromodomain; 1.
 DR Pfam: PF00628; PHD; 1.
 DR Pfam: PF00855; PWWP; 1.
 DR PRINTS: PR00503; BROMODOMAIN.
 DR SMART: SM00297; BROMO; 1.
 DR SMART: SM00249; PHD; 2.
 DR SMART: SM00293; PWWP; 1.
 DR PROSITE: PS00633; BROMODOMAIN_1; FALSE_NEG.
 DR PROSITE: PS50014; BROMODOMAIN_2; 1.
 DR PROSITE: PS50812; PWWP; 1.
 DR PROSITE: PS01359; ZF_PHD_1; 1.
 DR PROSITE: PS50016; ZF_PHD_2; 1.
 KM Zinc_finger; Bromodomain.
 FT ZN_FING 1 1
 FT ZN_FING 221 271 PHD-TYPE.
 FT DOMAIN 415 441 GLU-RICH.
 FT DOMAIN 615 685 BROMODOMAIN.
 FT DOMAIN 1085 1168 PWWP.
 FT SEQUENCE 1214 AA; 136598 MW; C6490810622109CD CRC64;

Query Match
 Best Local Similarity 3.9%; Score 105; DB 1; Length 1214;
 Matches 109; Conservative 65; Mismatches 181; Indels 130; Gaps 27;

QY 28 QSCOLIIEPPDEPNNGHTKSVSFREIYVLSHOVLNLYDILIEFYKSGSPGEEK 87
 DB 569 ERARLLIELIRKREKLRQGVQQAAMELELMPFNLLRTTLDLQE--KDPAHFAE 625
 QY 88 TIOVPEAKLAGFLYVISMQNAVIFDLDLS--YRTAREPPTSGCLLKKVSGIGAA 145
 DB 626 PVNLSE--VPDYLEFISKPMDEFSTMRKRLSHLRTLEEFED-----F 667
 QY 146 NLVRSQASFPN---IFYHALV-----CAVLTOETITAEQVKKVLFEDDERSTDSQ 195
 DB 668 NLIYTNCMKYNAKQITFHRAAVRLDGLGAILRH-----ARQAEINIGYDERCTHLPES 722
 QY 196 CSSFD-----ED-----IFEETQVSPRGKEKRRARAPLLSVQPVSNADWVLYKR 244
 DB 723 PKLEDYFRFSGMEDVDNLIIPENRAHLSPE--VQLKELLEKLDVLSARSSGA---RTRR 776
 QY 245 LHKRMELCNVYIOMHLDLNCMEEPPIFKGDPFPIIPSPQSSSTPSTG---GFSG-- 298
 DB 777 VRLRLREI--NALQKL-----AOPP-----PQPSLKVTVSNGELPAGPGDA 819
 QY 299 -----KETPSD--DRSQSR-----EHMGESLSLRAGGDLPLPSPKV-----EKK 338
 DB 820 AVLEQALQEEDEDDGDDSKLPPPTLEPGAPSL-----SEGESPEPEPTLKLINDSK 875
 QY 339 DPSR-----KKEWENA-----GNK-IYMAADKIYISKMTYKRRKQOHNSAFPE 385
 DB 876 PPSRFLPKRVEDELEKSPQLQGNELRLSLDNGINRLSLMAPDTPAGTPLSGVGR 935

OY 386 VKVEKKEPLGPGODSPLLOR-PQHLMDOGOMRHSFAGP-----ELLRODKRRPSG 437
DB 936 TSV-----LFKAKNGVKLQSRDPLVNGE-DHGAGSPASPAISIEEHRSKRRPSR 988
OY 438 STGSS-442
DB 989 SCSES 993

RESULT 13
VP41_CAEEL STANDARD: PRT: 876 AA.
ID VP41_CAEEL
AC Q19954;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vacuolar assembly protein VP541 homolog.
GN F32A6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Wu X.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
RP REVISIONS.
RA Material R.;
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY AND VACUOLAR TRAFFIC (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE VP541 FAMILY.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U40409; AAA81388.2;
DR WormPep: F32A6.3; CE29305.
DR InterPro: IPR000547; Clathrin_repeat.
DR InterPro: IPR004906; Transposase_Tc5.
DR InterPro: IPR001841; ZnfRing.
DR Pfam: PF00637; Clathrin_1; ZnfRing.
DR Pfam: PF03221; Transposase_Tc5; 1.
DR SMART: SM00299; CLH; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
DR Hypothetical protein; zinc-finger.
FT ZNFING 614 868
FT ZNFING 876 AA; 99987 MM; BD4BOE2AD4211C06 CRC64;
SQ SEQUENCE 876 AA; 99987 MM; BD4BOE2AD4211C06 CRC64;

Query Match 3.9%; Score 104.5; DB 1; Length 876;
Best Local Similarity 18.3%; Pred. No. 6.2;
Matches 66; Conservative 64; Mismatches 138; Indels 93; Gaps 15;

OY 1 RIRAMAQOVMLDTQCSPPK---TPNPDHAOSCOLIIEPDEKPNCHTKSKVSFRFETV 57
DB 425 RARAARLRPLIGCECKEEMWAVNGEEVLCITLAEVLDPDGP---TLDPECYQVLI 481
OY 58 SILSHOV-----LLQNLV-DILEEFV-----KGSPGEKTIQVEAKLAG 98
DB 482 ACLENNVKKQFRKLVQTVSPDLVYMSFTIDRTQWRIQIISGNIADVDTERVIMDALAH 541
OY 99 FLRY-----ISMQNLAV-----IFDLLDSYRTAFEDTSPGLKCLKLVKS 139

DB 542 LLYERKYESAKILMSCODEFOIFNVIDKHOLFVLVKQITELMNINSEALRLLLNAD 601
OY 140 GIGGA-----ANLYRSAMSFNIYFALVCAVLT-----NOETITAEQVKVFEEDERS-- 189
DB 602 SVSPSFYWEKIGRKQKTL-----AYITKLMSRNEGTEFPADKAVQIAETDQKLL 652
OY 190 -----TDSQOCSSEDEDIFFETAQVSPRKE-----KRWRRAPPLLSVQ 231
DB 653 PELRRNANYNVNKKARKLCS--DKGIIEETIYLLAKSGNHVDAVMKMYREYNMEKVIDYC 710
OY 232 PVSADWVVLVRLKILKLMELCNNTIOMHLDLNCMEPPPLFKODPFIILPFSQSESTP 291
DB 711 KDQNDPDLMI--HLGVAAEPFAHFSOLITIASNCL-----DPLIMDKLPDSDIP 760
OY 292 S 292
DB 761 N 761

RESULT 14
MRP6_RAT STANDARD: PRT: 1502 AA.
ID MRP6_RAT
AC O88269;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance-associated protein 6 (MRP-1-like protein-1)
DE (MLP-1).
GN ABCG6 OR MRP6 OR MLP1.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RX MEDLIN-98279126; PubMed-9614210;
RA Hitonashi T.; Suzuki H.; Ito K.; Ogawa K.; Kume K.; Shimizu T.;
RA Sugiyama Y.;
RT "Hepatic expression of multidrug resistance-associated protein-1-like
RT proteins maintained in e1sai hyperbilirubinemic rats.";
RL Mol. Pharmacol. 53:1068-1075(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RA Madon J.; Hagenbuch B.; Stieger B.; Meier P.J.;
RT "Cellular localization and functional characterization of a novel rat
RT liver multidrug resistance-associated protein (mrp6).";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
CC* -1- FUNCTION: MAY PARTICIPATE DIRECTLY IN THE ACTIVE TRANSPORT OF
CC DRUGS INTO SUBCELLULAR ORGANELLES OR INFLUENCE DRUG DISTRIBUTION
CC INDIRECTLY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- TISSUE SPECIFICITY: HIGH IN LIVER AND LOWER IN DUODENUM AND
CC KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB010466; BAA28954.1;
DR EMBL: U73038; AAD12747.1;
DR HSSP: P13569; INDB.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABCtransportTM.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00005; ABC_tran; 2.


```
FT DNA_BIND 2152 2211 HOMEBOX 1.
FT DNA_BIND 2249 2308 HOMEBOX 2.
FT ZN_FING 2335 2358 C2H2-TYPE (ATYPICAL).
FT ZN_FING 2539 2561 C2H2-TYPE.
FT DNA_BIND 2650 2709 HOMEBOX 3.
FT ZN_FING 2720 2743 C2H2-TYPE.
FT DNA_BIND 2952 3011 HOMEBOX 4.
FT ZN_FING 3032 3056 C2H2-TYPE.
FT ZN_FING 3552 3576 C2H2-TYPE.
FT ZN_FING 461 491 POLY-GLU.
FT DOMAIN 771 785 POLY-ALA.
FT DOMAIN 1314 1317 POLY-ALA.
FT DOMAIN 1734 1748 POLY-GLN.
FT DOMAIN 1794 1799 POLY-GLN.
FT DOMAIN 1856 1863 POLY-GLN.
FT DOMAIN 2044 2059 POLY-PRO.
FT DOMAIN 2405 2408 POLY-ALA.
FT DOMAIN 3216 3220 POLY-PRO.
FT DOMAIN 3380 3409 POLY-GLN.
FT DOMAIN 3412 3420 POLY-GLN.
FT DOMAIN 3534 3550 POLY-GLY.
FT DOMAIN 3620 3623 POLY-PRO.
FT DOMAIN 3659 3662 POLY-SER.
SQ SEQUENCE 3726 AA; 406567 MM; 915ACBE588A72C98 CRC64;
```

Query Match 3.88; Score 103; DB 1; Length 3726;
Best Local Similarity 19.7%; Pred. No. 56;
Matches 72; Conservative 49; Mismatches 86; Indels 158; Gaps 19;

```
QY 132 KCLLKVKYSGIGGANLYRQASMSNIYFHLVCAVL-----NOETIT 174
DB 1073 KHLQEHESGVEGE-----SCYHCVLCTYSTKAKLNLQHVRSKMHORSESLRK 1121
QY 175 AEQVKVLFEDDE-----RSTDSQ-----QCSSEDEDFEETQAVSPPR 214
DB 1122 LQRLQKGLPEDEDLQIFITIRCPSTDPPEPYEDAGPSPASADPELAKDQSGS-EE 1180
QY 215 GKERR-----QNRARPLLSVOP-----V 233
DB 1181 GQSKRAASSQAERELTDSPATYTKRTSFPGSSSETPLSSKRPKASEEIKPEOMYOCPCYKY 1240
QY 234 SNADWVLYVRLH-----KICMEICNN--YIOMHLD-----LNCMEPEPP 271
DB 1241 SNAD---VNRRLVHANTQHSVQPLRLRCPLCQDKMLNKKIHQLHLTHLSVAPDCVEKLI 1296
QY 272 IFKGDPEFFILPS--FOSESSTPTSGFS---GKETPSEDDRSQ-----SREHMGESLS 319
DB 1297 MVTAPPEMVPSSMFLPAAADNRDGNSTLEEVGKQPEASEDPGKNILPSPASMEH----- 1350
QY 320 LKAGGDDLPPSPKVEKKDPSRKKE-----WMENAGNKIYTMADKTIKMLMEYK--- 371
DB 1351 ---GGDL-----KPTSDPSCGREDSGFLCWKKGCONQYF-----KTSATLQTYHNEVH 1395
QY 372 -KRKQ 375
DB 1396 AKRPQ 1400
```

Search completed: June 17, 2003, 11:56:44
Job time : 34.0952 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:36 ; Search time 55.3646 Seconds
(without alignments)
1000.321 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRMAQGVFLMDTQCSPKT.....VRAVREWLGRVRYDITV 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	4.1	828	US-08-681-219-28	Sequence 28, Appl
2	108.5	4.0	2783	US-09-816-669A-14	Sequence 14, Appl
3	100.5	3.7	1294	US-09-836-499-2	Sequence 2, Appl
4	100.5	3.7	1294	US-10-162-435-2	Sequence 2, Appl
5	100.5	3.7	1309	US-09-836-499-5	Sequence 5, Appl
6	100.5	3.7	1309	US-10-162-435-5	Sequence 5, Appl
7	98	3.6	411	US-09-768-600-5	Sequence 5, Appl
8	98	3.6	446	US-09-738-626-4700	Sequence 4700, Ap
9	98	3.6	1332	US-09-982-091A-4	Sequence 4, Appl
10	97.5	3.6	669	US-09-823-187-88	Sequence 88, Appl
11	96.5	3.6	759	US-09-764-864-823	Sequence 823, App
12	96	3.6	555	US-09-764-864-1161	Sequence 1161, Ap
13	96	3.6	555	US-09-764-864-1574	Sequence 1574, Ap
14	95.5	3.5	693	US-10-151-569-2	Sequence 2, Appl
15	95.5	3.5	1713	US-10-171-311-113	Sequence 113, Appl
16	95	3.5	678	US-09-823-187-87	Sequence 87, Appl
17	95	3.5	1395	US-10-153-668-468	Sequence 468, App
18	94.5	3.5	465	US-10-103-313-357	Sequence 357, App
19	94.5	3.5	504	US-09-801-368-208	Sequence 208, App

20	94.5	3.5	600	US-09-764-864-1282	Sequence 1282, Ap
21	94.5	3.5	907	US-10-200-154-4	Sequence 4, Appl
22	94.5	3.5	907	US-09-954-043-4	Sequence 4, Appl
23	94.5	3.5	940	US-10-102-806-678	Sequence 678, App
24	94.5	3.5	1711	US-09-771-161A-219	Sequence 219, App
25	94.5	3.5	1711	US-09-771-161A-220	Sequence 220, App
26	94	3.5	258	US-09-815-242-5680	Sequence 5680, Ap
27	94	3.5	300	US-09-815-242-12173	Sequence 12173, A
28	93.5	3.5	324	US-09-987-107-7	Sequence 7, Appl
29	93.5	3.5	323	US-09-987-107-58	Sequence 58, Appl
30	93	3.4	266	US-09-864-761-37015	Sequence 37015, A
31	93	3.4	464	US-09-902-941-1934	Sequence 1934, Ap
32	93	3.4	464	US-10-017-754-1934	Sequence 1934, Ap
33	93	3.4	615	US-09-925-301-1094	Sequence 1094, Ap
34	93	3.4	693	US-10-029-217A-4	Sequence 4, Appl
35	93	3.4	705	US-10-154-386-2	Sequence 2, Appl
36	93	3.4	1038	US-09-908-500A-2	Sequence 2, Appl
37	93	3.4	4019	US-09-854-133-425	Sequence 425, App
38	93	3.4	4019	US-09-738-973-425	Sequence 425, App
39	92.5	3.4	393	US-09-745-763-19	Sequence 19, Appl
40	92.5	3.4	928	US-09-801-368-282	Sequence 282, App
41	92.5	3.4	1080	US-09-904-380-2	Sequence 2, Appl
42	92.5	3.4	2843	US-08-681-219-32	Sequence 32, Appl
43	92	3.4	397	US-09-821-883-27	Sequence 27, Appl
44	91.5	3.4	304	US-09-987-107-6	Sequence 6, Appl
45	91.5	3.4	323	US-09-987-107-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-08-681-219-28
Sequence 28, Application US/08681219
Patent No. US20020058607A1
GENERAL INFORMATION:
APPLICANT: Takaaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USBS THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48963/JFW/JKM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO.: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 828 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-681-219-28

Query Match 4.1%; Score 110; DB 8; Length 828;
 Best Local Similarity 18.4%; Pred. No. 0.81;
 Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

```

OY 81 PSPGEKTI-----QVPEAKGLAFRLYRIMONLAVIFDILLDSRYRAREDRSPGLKCLK 136
DB 306 PSTGLSTLSSSSNDIPKIKIERYK-----LSTRESSSSSDPRVIGS 348
OY 137 KVSIGIGANLYROSAMSF-----NIYHALVCALVTNOETTTAEQVKLFED 185
DB 349 EISSIGVSSVAEHLAHSLODCSNIOEIFQTLYSHG-----SAISEKIRE--FEV 397
OY 186 DERSDSSQCCSDEDEDFEETAQVSPRGKEKRWARMPLLSQVPSNADWVLVRL 245
DB 398 ETERLSRIEHLKSONDLITLLE-----ECKSNABRMSMLVGYKESNATRLAQL 450
OY 246 HKLCELCNNYIOMHLDLNCMEEPPIKGDPEFILPFSQESSPTSGSGKETPSE- 304
DB 451 SEQCEAEVELL-----ALAESQSLILGFPRAAGVSSFGDOSGDENITOM 497
OY 305 ----DDRQSREHMGESLSLACG--GDLLPSPSKVEKKDPSRKKEWENAGNIYMA 358
DB 498 LKRAHDCRKTAEKNAKALMLKLDSCGGAFAVAGCSVP-----WESLSSNSHST 548
OY 359 ADKTSKLTETEKRRKROOHNSAPPEKVEYVEKKG-----EPLG----PRGD 401
DB 549 TSSSTASSCDTEFTKEDEO-RLKDYIQOLKNDRAVAKLMLLESHIDPLSTYDVKPRDS 607
OY 402 SPL-LQRP--OHLMDQOMRHSFAPPELLRODKRPRSSTGSSLSVSDAEAOIQAW 457
DB 608 QRLDENAVLMOELMAKEMAEKLAQLYLEKEKK-----ALEIKLSTREAOEQAY 659
OY 458 TNNVLTVINOIQLPDDQFTALQ-----PAYFPCS-----QUTCHVTDIR 498
DB 660 LVHIEHLSEVEQEKQOMRSLSTSSGSKDKPGKECADASPALSLAELRTTSENELA 719
OY 499 VR--QAVREWLGRVGRVYDII 517
DB 720 AEFNAIRREKKLAKARQELV 740

```

RESULT 2

US-09-816-669A-14
 ; Sequence 14, Application US/09816669A
 ; Patent No. US20020137019A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GARABEDIAN, Michael
 ; APPLICANT: TANEJA, Samir
 ; APPLICANT: HITTELMAN, Adam
 ; APPLICANT: MARKUS, Steven
 ; TITLE OF INVENTION: METHOD FOR SCREENING TRANSCRIPTIONAL COREGULATORY PROTEINS OF
 ; TITLE OF INVENTION: TRANSCRIPTION FACTORS, AND ANDROGEN RECEPTOR-DEPENDENT DISEASES
 ; TITLE OF INVENTION: PROTEINS AS TARGETS FOR ANDROGEN RECEPTOR-DEPENDENT DISEASES
 ; FILE REFERENCE: GARABEDIAN-1.1A
 ; CURRENT APPLICATION NUMBER: US/09/816, 669A
 ; PRIOR FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/225, 618
 ; PRIOR FILING DATE: 2000-08-15
 ; PRIOR APPLICATION NUMBER: 60/191, 768
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 14
 ; LENGTH: 2783
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-816-669A-14

Query Match 4.0%; Score 108.5; DB 10; Length 2783;
 Best Local Similarity 20.6%; Pred. No. 6.5;
 Matches 73; Conservative 45; Mismatches 91; Indels 145; Gaps 18;
 OY 132 KCLLKRVSGIGGANLYROSAMSFNIYFHALVCALVT-----NQETIT 174

```

DB 158 KHLQOHESGVEGE-----SCYHCVLCYNSTRAKINLIQHVRSKMHQRESLRK 206
OY 175 AEOVKVLFEEDE-----RSTD-----SSQ 194
DB 207 LORLKGLEPEDEDEGLQFTTRRCPSIDPEEALIEDVEBPSETADPEELANDQEGCASS 266
OY 195 QCSSEDEDFEETAQVSPRGKEKRWARMPLLSVQP-----VSNA 236
DB 267 QAEKELTDSPATSKRISFPGSES-----PLSKRKRTABEIKPEQMOCPYCKNSA 319
OY 237 DMWLVYKRLH-----KLCMELCNN--YIOMHLD-----LENCMEEPLTFK 274
DB 320 D-----VNRLRYHAMTQHSVQPMRLRCPLQODMLNNKHILQHLTHLSVAPDCVEKRLMTV 375
OY 275 GDPFPLPSFQ--SESTSTGSGFSGKETPSEDDRQSRHMGESLSLACG--GGDLLP 330
DB 376 TTPEKVMSSMPLRAVADRDGNSNLEBAGK--QETISEDGLKNTLPSASTEQSGD--LK 431
OY 331 PSPKVEKKDPSRKE-----WENAGNIYMAADKTSKLTETEK-----RRKQ 375
DB 432 PSP-----ADPGSVREDSGFICWKKCQNVF-----KTSALQTHFNEVHAKRPQ 476

```

RESULT 3

US-09-836-499-2
 ; Sequence 2, Application US/09836499
 ; Publication No. US20030027316A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel
 ; APPLICANT: AND 16051B, NOVEL HUMAN PDZ
 ; TITLE OF INVENTION: FAMILY MEMBERS AND USES THEREOF
 ; FILE REFERENCE: 10448-043001
 ; CURRENT APPLICATION NUMBER: US/09/836, 499
 ; PRIOR FILING DATE: 2001-04-17
 ; PRIOR APPLICATION NUMBER: US 60/197, 507
 ; PRIOR FILING DATE: 2000-04-18
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1294
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-836-499-2

Query Match 3.7%; Score 100.5; DB 9; Length 1294;
 Best Local Similarity 21.8%; Pred. No. 11;
 Matches 111; Conservative 66; Mismatches 194; Indels 139; Gaps 24;

```

OY 10 FMLDTQCSPTKPPNNDHAGSOLITIELP-----PDEKPMGHTKKYSFREIYV----- 57
DB 839 FNMVAIRMIONSPDNI-----ELIISQSKGVGNPDDEKNGTANSGVSTDLISFGYQG 892
OY 58 SLTSH-----QVLLQNLVYDILLEFVAG-----PSPGEKTIQVPEAK 95
DB 893 SLSTHTQDQDNTELMAGVQSLVPRLRHQLSLFLPKGAASSCPSPPELSAGEITFVE 952
OY 96 LA---GFLRY-----ISMQNLAVIFDILLDSTRTAREFTPSGLCKLKRVSGIGAA 145
DB 953 LKEDGTGLGFSVTGINTSVPGYGYKSIYVPGSPAAGEGIILOGDRL--QVNGV---- 1006
OY 146 NLVQGSAMSFNIYFHALVCALVTNOETTTA-----EQYKVLFEEDERSTDSQCCSSPDE 201
DB 1007 -----ILCG--LTHQAVQCLKGGQVAVRLVLE--RRVPRSTQCCPSAND 1047
OY 202 DIFETQAVSPRGKEKRWARMPLLSVQP-----VSNADWVLVYKRLKLMELCNN 255
DB 1048 SMGDERTRAVS-----LVYALPGRSSCVSVTDGKRFKFKLKKANGLGFS 1092
OY 256 YIOMHLDLNCMEEPPIKGDPEFILPFSQESSPTSGSGKETPSEDDRQSRHMG 315
DB 1093 FVQM--EKESCSH-----LKSDLVRIKRLFLPGQPAENCAIAGDIIILAVNGRSTEGILFQ 1146

```

```

QY      316  ESLSLKAGGGD---LLPSPRYKKEDKPSKKKEW---ENAGNKITYMA--ADTITSLM 367
D      1147  EVTLHLGLGAQEVTLILLCRPPGAL--PEMEQWMPHETLSADKEFTFRATCTDCTSPIL 1203
QY      368  TEYKRRKQNHLSAFPKEVKYKKKGEPGPGGDSPLLGRPQHMLDGGOMR----- 418
D      1204  DQEDSWMD---SASP-----DAGEGIGLEPRESS--QKAIREAQWGNRRERPAWSSLT 1256
QY      419  HSFAGPELLRQDKRRSGSGSSISYSVR 448
D      1251  HSPESHPLCKLQHERDESTLATYLEKDYR 1280

```

```

RESULT 4
US-10-162-435-2
; Sequence 2, Application US/101612435
; Publication No. US20030096305A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Glucksman, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Kappeller-Liberman, Rosana
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
; TITLE OF INVENTION: CELL SURFACE PROTEIN FAMILY MEMBERS
; FILE REFERENCE: 10448-189001
; CURRENT APPLICATION NUMBER: US/10/162,435
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US01/12420
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/891,008
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19963
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,220
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/860,868
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16013
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,674
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/886,429
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/20055
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,963
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: PCT/US02/00275
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/260,286
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/41811
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,612
; PRIOR FILING DATE: 2000-08-21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-435-2

```

Query Match	3.7%	Score 100.5;	DB 9;	Length 1294;
Best Local Similarity	21.8%	Pred. No. 11;		
Matches 111; Conservative	66;	Mismatches 194;	Indels 139;	Gaps 24;

```

QY. 10 FMDTQSPKRNPHNHAQSCOLLITELP-----PDKRPMGHKTKSFEIIV----- 57
Db 839 FNNAAVRIQNSPDNI-----ELLISQSKVGGNPNDEEKKNGTANSVSSDILSFGYQ 892
QY 58 SLUSH-----QVLLQNLXYDILLEEYVG-----PSDEKRTIQVPEAK 95
Db 893 SILSHTDQDDBNTEELDMAGVQSIVPRLRHOLSEPLTKGAGSSCPSPPEISAGEIYEVE 952
QY 96 LA---GLRLY-----ISMQNLAVIPDILLDSYRTAEFPSTPGLKLLKVGSGIGAA 145
Db 953 LYKEDGTLGFSVTGCGINTSVYGGIYKASIVPGCPAAEGGIILOQDRLL--QVDBV---- 1006
QY 146 NLRYQASMFNIYFHALVCAVLITQETITA---EYQKVLFEEDERSTDSQOCCSEDE 201
Db 1007 -----ILCG-LTHKQAVQCLKGQVAVRLVLE--RVRPRSTQCCPSAND 1047
QY 202 DIFETRYQVSPPRGKREKQMARAPLISVQF-----VSANDWIVLYRLKIKELCINN 255
Db 1048 SMGERTAVS-----LVYALGPRSSCVSYVDGPKFEYKLLKNANGLGFS 1094
QY 256 YIQHLDLENCEMEERPIFGDPFEILPSFOGESSPTSYGFSGKETPSEDDRQSGREHMG 315
Db 1093 FVQW--EKESGSH---LKSDLVRIKRLFPQCPRAENCAIAGDIIILAVNGSTGLFFO 1146
QY 316 ESLSLAKGAGD---LLPSPRYEKKDPKSRKKEWM--ENAGNKITYMA--ADRTISKLM 367
Db 1147 EYVLLHLGAPQEVYTLTLCRPPGAL--PEMEQEQNPBELSADKEFTATCTDCTSPIL 1203
QY 368 TEYKRRKQOHNLSAFPKREVKEKGEPLGPGGODSPILQORONHLDQOQMR----- 418
Db 1204 DQEDSWRD---SASP-----DAGEGLGLRPRESS--QKAIREAMQONRRERPAWSSUT 1256
QY 419 HSFSAQPELLQDKRRPGRSGTSSISVSVR 448
Db 1251 HSPESHPIELCLHQRDESTLATSLEKDYR 1280

```

RESULT 5
US-09-836-499-5
Sequence 5, Application US/09836499
Publication No. US20030027316A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 16051A AND 16051B, NOVEL HUMAN PDZ
FILE REFERENCE: 10448-043001
CURRENT APPLICATION NUMBER: US/09/836.499
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: US 60/197,507
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ. ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1309
TYPE: PRT
ORGANISM: Homo sapiens
US-09-836-499-5

	Query Match	3.7%	Score 100.5;	DB 9;	length 1309;	
	Best Local Similarity	21.8%	Pred. No.	11;		
	Matches	111:	Conservative	66;	Mismatches	194; Indels 139; Gaps 24
QY	10 FMLTQCSPKPPNPNFNAHSCQLLIELP-----PDKEPNCHTKYSVFRILIV-----	57				
DG	I- - - - -I-I-	- - - - -E-LIIISQSGVGNGNPDEEKINGSTANGSVSDTLISFGYGOG	892			
	839 FNMAVRMIQNPSDNI-----ELTISSQSGVGNGNPDEEKINGSTANGSVSDTLISFGYGOG	892				
QY	58 SLISH-----QVLNLNQYLIDLLLEEFVK-----PSGEKKITGVPEAK	95				
bB	893 SILSHTDODADNTAEELDAGVOAIVPRLRHQLSFLPLRIGACGGSCPPELTSAGITVE	952				

```

OY      96 LA-----GFLRY-----ISMONTAVIFEDLDLSYRTAREFDPTSGLKCLKLKVSIGGAA 144
Db      953 LYKEDGTLGFSVYTGCIITSVYGGIITYKSIYPGCPAKKEGIILOGDRLL--QYDGV----- 1006
OY      146 NLYROSAMSFNIYFHALVCAYLVNQETTTA---EYQKYLFEEDERSTDSQCCSEDE 201
Db      1007 -----ILTCG-LTHKQAVQCLKGQVARYLVE--RRVPRSTQCPSAND 1047
OY      202 DIFEEATQVSPRBEKKEKRWARARPLLSVQP-----VSANDWVWLKRLHKLCLMELCNN 255
Db      1048 SMGDERTAVS-----LVTALGRPSSCVSYVDGKRFVEYKLLKNANGIGFS 1092
OY      256 YIQMHLDLENCEMEEPPIFKDPFFILTSFQSESTPSTGSGFETPSEDDRSOSREHM 315
Db      1093 FVQM--EKESCSH-----LKSGLVYIKRFLFPQGPAPENALIAAGIILAVNRSYIEGLIFQ 1146
OY      316 ESLSLKAGGD---LLIPSPKYKKKDPSSKKKEW--ENAGNKIYTMA--ADTKTKLM 367
Db      1147 EVLLHLKRAPQEVYLLILCRPPGAL--PEMEQEWQPELISADKEFTTRATCTDCTSPIL 1203
OY      368 TEYKRRKQNHLSAFPEKVEKYKKEGPELPGGQSPILQRRQHLMDOGMR----- 418
Db      1204 DQEDSWRP---SASP-----DAGEGIGLPESS--QKAIREAQWQNRERPWASSLT 1250
OY      419 HSFSGPELLRQDKRPRSSTGSSLSYSVR 448
Db      1251 HSPESHPLCLKLHQERDESTLATISLEKQVR 1280

RESULT 6
US-10-162-435-5
: Sequence 5, Application US/10162435
: Publication No. US20030096305a1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel
: APPLICANT: Glucksman, Maria Alexandra
: APPLICANT: Curtis, Rory A. J.
: APPLICANT: Kapeller-Libermann, Rosana
: APPLICANT: Bandaru, Rajasekhar
: APPLICANT: leihey, Kevin R.
: TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
: FILE REFERENCE: CELL SURFACE PROTEIN FAMILY MEMBERS
: CURRENT APPLICATION NUMBER: US/10/162,435
: CURRENT FILING DATE: 2002-06-04
: PRIOR APPLICATION NUMBER: US 09/836,499
: PRIOR FILING DATE: 2001-04-17
: PRIOR APPLICATION NUMBER: PCT/US01/12420
: PRIOR FILING DATE: 2001-04-17
: PRIOR APPLICATION NUMBER: US 60/197,507
: PRIOR FILING DATE: 2000-04-18
: PRIOR APPLICATION NUMBER: US 09/891,008
: PRIOR FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: PCT/US01/19963
: PRIOR FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: US 60/214,220
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 09/860,868
: PRIOR FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: PCT/US01/16013
: PRIOR FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: US 60/205,674
: PRIOR FILING DATE: 2000-05-19
: PRIOR APPLICATION NUMBER: US 09/886,429
: PRIOR FILING DATE: 2001-06-21
: PRIOR APPLICATION NUMBER: PCT/US01/20055
: PRIOR FILING DATE: 2001-06-21
: PRIOR APPLICATION NUMBER: US 60/213,963
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 10/041,406
: PRIOR FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: PCT/US02/00275

```

```

      PRIOR FILLING DATE: 2002-01-08
      PRIOR APPLICATION NUMBER: US 60/260,286
      PRIOR FILLING DATE: 2001-01-08
      PRIOR APPLICATION NUMBER: US 09/934,268
      PRIOR FILLING DATE: 2001-08-21
      PRIOR APPLICATION NUMBER: PCT/US01/41811
      PRIOR FILLING DATE: 2001-08-21
      PRIOR APPLICATION NUMBER: US 60/226,612
      PRIOR FILLING DATE: 2000-08-21
      NUMBER OF SEQ ID NOS: 38
      SOFTWARE: FASTSEQ for Windows Version 4.0
      SEQ ID NO 5
      LENGTH: 1309
      TYPE: PRT
      ORGANISM: Homo sapiens
      US-10-162-435-5

      Query Match      3.7%; Score 100.5; DB 9; Length 1309;
      Best Local Similarity 21.8%; Pred. No. 11;
      Matches 111; Conservative 66; Mismatches 194; Indels 139; Gaps 24;

      QY      10 FMLDTQCSPPKPNNEHDHAAOCLILELP-----PDEKRNCHTKKVSFRETIV----- 57
      Db      839 FNNAVRMKIQNSEPDNI-----ELIISQKGVGCGNNPDEKKNNTANGSVSTIILSGCYG 892

      QY      58 SLLSH-----QVLLQNLVYLILLEPVGK-----PSPEEKTIOYPEAK 95
      Db      893 SLLSHTQDQDRNTEELDMAGVOQLVPRRLHQLSFLPKAGSSCPSPREISAGELTFVE 952

      QY      96 LA---GGLRY-----ISMQNLAVIFDLILDSYTRAREFDTSPGLKCLIKKVGIGGAA 145
      Db      953 LKEDGTIGFVSYTGGINTSVPRGGIYKVSIVYGGRAKGGQLQDGRLL--QYDGY----- 1006

      QY      146 NLYRQASAMFNIYFHALVCALFTNQETTA---EQVKVLEFEDDERSTDSQSCSEDE 201
      Db      1007 -----ILCG-LTHKQAVOCLKPGQAVALE--RRVPRSTQCCPSAND 1047

      QY      202 DIFETIAGVSPRRCKEKQWRARMPLLSVQ-----VSNADWVWLYVKRLIKMLCMLQNN 255
      Db      1048 SMGDERIVAS-----LVITALPERPSSCVSVTDGPKREYVLLKKNANGLGFS 1092

      QY      256 YIQMHLIDNCEMEERPIEFKGPFFILPSQSESSPTSTGSGKETPSEDDRSQSREHMG 315
      Db      1093 FVQW--EKESCSH---LKSDIVRIKRLPRGQPAENGALIAAGDIIILAVNGRTEGLIRQ 1146

      QY      316 ESLSLKAGGCD---LLPRSPKVEKKDPSRKKEWM--ENAGKIKITMA--ADKTSIKLM 367
      Db      1147 EVLHLIRLAPQREVLTLLCRPPGAL--PEMQEWMQTPELASDKETFRATCTDSCSPIL 1203

      QY      368 TEYKKRKQOHNLSAPRKVEKKKGEPRGPRQDPSLILRQNLHMOGGM----- 418
      Db      1204 DQEDSWRD---SASP-----DAGEGLGLPRESS---OKAIREADWGQNNRREPPASSILT 1250

      QY      419 HSFSAGPELLRODKRRRSGSTGSSLSVSVR 448
      Db      1251 HSPESHPIKLHQERDESTLATISLEKQVR 1280

      RESULT 7
      US-09-788-600-5
      Sequence 5, Application US/09788600
      Patent No. US20020004489A1
      GENERAL INFORMATION:
      APPLICANT: Shi et al.
      TITLE OF INVENTION: Retinoid Receptor Interacting Polynucleotides, Polypeptides,
      TITLE OF INVENTION: Antibodies
      FILE REFERENCE: PTO17PI
      CURRENT APPLICATION NUMBER: US/09/788,600
      CURRENT FILING DATE: 2001-02-23
      PRIOR APPLICATION NUMBER: PCT/US00/22351
      PRIOR FILING DATE: 2000-08-15
      PRIOR APPLICATION NUMBER: 60/189,026
      PRIOR FILING DATE: 2000-03-14

```

;; PRIOR APPLICATION NUMBER: 60/148,757
 ;; PRIOR FILING DATE: 1999-08-16
 ;; NUMBER OF SEQ ID NOS: 7
 ;; SOFTWARE: PatentIn Ver. 2.0
 ;; SEQ ID NO 5
 ;; LENGTH: 411
 ;; TYPE: PRT
 ;; ORGANISM: Homo sapiens
 US-09-788-600-5

Query Match 3.6%; Score 98; DB 10; Length 411;
 Best Local Similarity 19.1%; Pred. No. 3.2;
 Matches 74; Conservative 54; Mismatches 115; Indels 144; Gaps 17;

```

QY 2 IRMAAOVFMLDQCSPKPNPNDHQAOSCOLIELEPPEKNGHTKKSVSFRETIVSLLS 61
DB 73 IKSLEKEELLE---EPTT---SHGSSQCIIVETSEE-----GNSVPASQVAULTS 119
QY 62 HOVLQNLXLDLLEFVKGSPSGEKTIOVPEAKLAGFLRYISMGNLAVIFDLIDSYRT 121
DB 120 KRSI-----VLME-----SSAEITY-CPETQLSS-----SEFDL----- 150
QY 122 AREDTSPGLCKLKKV-----SGIGCANLYROSAMS-----FN 156
DB 151 --EKEVSPGSHDLIDGVRIIMADREVGKKEDEKVAISTESSNOVSCPLDCQFPPTK 208
QY 157 IYFHALVCAVLTNOETITAEQVKVLFEDERSTDSQCSSEDEDFEETAQVSPRGK 216
DB 209 IERRAMCNGMEDYVITRQKEAKTKSDSGT-----AAQTSILDDK 251
QY 217 EKROWRAMPLLSVQVPVSNADWMLVKRLKLCMELCNNTI-----QMHLENCMEERP 271
DB 252 NEKCY-----LCKSLVFPFREYQCHVD--SCLQIAK 279
QY 272 IFKGDPEFLPSFQSESTPTSGGFSCK-----ETPSDDRSQSR-----EIMGSL 318
DB 280 ADQGD-----GPEGSGRACSTVEGKMQORLNPKKGSHSEGLLSFLDQSEHKTSDA 331
QY 319 SLKAGGDDLPLPSPKVEKKKDPSSRKE 345
DB 332 DIKSETGAFRVPSPGMEAGCSRMQ 358

```

RESULT 8
 US-09-738-626-4700
 ; Sequence 4700, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIRO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SEMOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4700
 ; LENGTH: 446
 ; TYPE: PRT

;; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4700

Query Match 3.6%; Score 98; DB 9; Length 446;
 Best Local Similarity 21.1%; Pred. No. 3.7;
 Matches 96; Conservative 59; Mismatches 149; Indels 152; Gaps 23;

```

QY 125 FDTSPGLCKLKKVSGIGCANLYROSAMSFNIFHALVAVL-----TNOETIAE 176
DB 31 FEIAPGERILILTAGSACKSTLL-----AALAGVLGSDGSGVSTGELLVAP 77
QY 177 QVKVLFEDDER-----STDSSQCSS-----EDEDIFETAQ-----VSPPR--- 214
DB 78 SIGLVLPDQSDVIASTRIGDVAFCEMQLQIPREIWRVREALVELGDLPLSHPTKYL 137
QY 215 --GKKR-----QWRAMPILLSVQVPVSNAD-----WMLVKRL 245
DB 138 SGGOKORALAGVIMAGARLILLD--EPTANLDPOGOKDVAAVDRVOETGATLLIVEHR 196
QY 246 HKLCMELCNNTIOMHLENCMEERPPIKGPDFILPSFQSESTPT-----STGGF 296
DB 197 HELWNIITDRIST--TDGEDVOPAEILIVGQ-----LPGAQSTSKPILMANDLCTWGL 251
QY 297 SKETPSEDDRSQSRHNG-----ESLSKAGGDDLPLPSPKVEKKD-----PSR 342
DB 252 RSFEVP-----EGASTVITGPNAGKSTLALTWGG--LLPCKSQLELSDTVRGGLNTPPH 305
QY 343 KKEWEN--GNKITMAADKTIKLMTEYKKRKOHLNLAFFPKVYKKEKGPPLPGQ 400
DB 306 K--WRSADLAARIGTVQDPE--HQAFAVTRDELEIGPKIMKVDAS----- 348
QY 401 DSPLLRQPHLMDQOMRHSFSAGPELLROKRPSSGTSLSVSVDAEAOIQAMTNM 460
DB 349 -----ERIEELDLRLRLHLENANFTL-----SGGEKRLSVAT----- 383
QY 461 VLVNLNQIQLIPDOTFTALQPAVPPCISQLTCHVTD 496
DB 384 ALVAPKLLILDEPTF--GQDEPTETELVTMLRELTD 418

```

RESULT 9
 US-09-982-091A-4
 ; Sequence 4, Application US/09982091A
 ; Patent No. US20020151030A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
 ; APPLICANT: KUMAGAI, AKIO
 ; APPLICANT: DUNPHY, WILLIAM
 ; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
 ; FILE REFERENCE: CIT1320-1
 ; CURRENT APPLICATION NUMBER: US/09/982,091A
 ; CURRENT FILING DATE: 2002-10-17
 ; PRIOR APPLICATION NUMBER: US 60/241,246
 ; PRIOR FILING DATE: 2000-10-17
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1332
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-982-091A-4

Query Match 3.6%; Score 98; DB 10; Length 1332;
 Best Local Similarity 20.4%; Pred. No. 18;
 Matches 73; Conservative 53; Mismatches 124; Indels 108; Gaps 15;

```

QY 168 TNOETITAEQV--KVFEDERSTDSQCSSEDEDFEETAQVSPRGKPKROWRAM 225
DB 46 SDEIFVSKIKLNKRVLODSDSETDTWASPEKTYDAAEENKENLYAGKNTK----- 99
QY 226 PLSVQVPVSNADWMLVKRLKLCMELCNNTIOMHLENC-----CMEEPRIRKGPDP 278
DB 100 -----IKRIKTVADSDSESTYKSLYOENLEAOVKPCLE----- 133

```

	Query March	3.6%	Score 97.5;	DB 9,	length 669;
	Best Local Similarity	18.6%;	Pred. No. 7,4';		
	Matches	79;	Conservative	57; Mismatches 123;	Indels 163; Gaps 18'
OY	184 EDDERSTDSQCCSEDEDIREFEAQAQSPP-----GKEKROWRAMPPLSYQPYSNADW	238			
	: : : : : : : : : : : : : : : : : : :				
Dd	186 EEDESSESSESKTSDDOTPEPKKA-ARPRGLPGKKKK-----VPASDSDS-	234			
	: : : : : : : : : : : : : : : : : :				

```

1 SEQ ID NO 823
2 LENGTH: 759
3 TYPE: PRT
4 ORGANISM: Homo sapiens
5 FEATURE:
6 NAME/KEY: SITE
7 LOCATION: (18)
8 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
9 NAME/KEY: SITE
10 LOCATION: (19)
11 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
12 NAME/KEY: SITE
13 LOCATION: (21)
14 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
15 NAME/KEY: SITE
16 LOCATION: (257)
17 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
18 NAME/KEY: SITE
19 LOCATION: (299)
20 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
21 OS-09-764-864-823

```

Query Match	3.6%	Score 96.5	DB 10	length 759
Best Local Similarity	19.1%	Pred. No. 11		
Matches	101	Conservative	84	Mismatches 184; Indels 159; Gaps 26
QY	40	EKPNGHTKKSYSFR-----	ELVSLSLHQVLTQMLYDILLEEFKGGSP	83
Db	290	QKAMDHIEEXSLIKCKCKAAELVATYHFSHGIEIVIKKIQVLYKFLKSLD----	PRE	344
QY	84	G---EETKIQQPEAKLQGLFRLYRISMQLAVITDLL--LDYSR-----	TAREEDTSPGK	132


```

RESULT 14
US-10-151-569-2
? Sequence 2, Application US/10151569
? Publication No. US20030040050A1
? GENERAL INFORMATION:
? APPLICANT: Matsumoto, Kunihiko
? TITLE OF INVENTION: NOVEL PROTEIN TAB2
? FILE REFERENCE: 06501-110US1
? CURRENT APPLICATION NUMBER: US/10/151,569
? CURRENT FILING DATE: 2002-05-20
? PRIOR APPLICATION NUMBER: PCT/JP99/06466
? PRIOR FILING DATE: 1999-11-19
? NUMBER OF SEQ. ID. NOS: 19
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 2
? LENGTH: 693
? TYPE: PRT
? ORGANISM: Homo sapiens

```

US-10-151-569-2

Query Match 3.5%, Score 95.5; DB 9; Length 693;
Best Local Similarity 18.1%, Pred. No. 12;
Matches 102; Conservative 76; Mismatches 182; Indels 203; Gaps 24;

QY 61 SHQVLLONLVDILLEEFYKGPSPGEKTIQVPEAKLAGELRYISMQ-----LAVIED 113
DB 5 SHQIDFQVLDL-----ROKFEVEPVVVSRCM-----LQNNNNLACCAYL-- 46
QY 114 LLDSYRFR-----EFDTPGLCKLKKVSGIGAANLYROSANSMFYTHALVCA 165
DB 47 -----SQESTRLYIEGDLNESDSDGIGSLRNHMTSL-----NLDIQSQ-----NITHGREGS 95
QY 166 VLTNOETTTAQQVKKVLFEDDERSTDSQOCSEDEDFEETAOVSPPRGKEKROWRAM 225
DB 96 RMNGSRRLT-----HSISDQIQGQSGNSELFQOEPOQAP-----AQV 133
QY 226 P-----LSTQPVSNADWVWLKRLKLCMELCNNTYIOMHL-----DLNCEMEEPPIFKG 275
DB 134 POGFNVFGMSSSGASNS-----APHLGFHLGSKGTSSLQOTPRPNP 176
QY 276 DPFILPSFQESSSTPTSGFGSKETP-----SEDDRS-----QSRHMGES 317
DB 177 IMVTLANIQGRNTPTSLHIGVPPVPLNSPOGNSIYIRYITTPGTGTQTOQHSQWV 236
QY 318 LSLKAGGDDLPPSPKVEKKDPSRKKEWE-----NAGNKITYMA 358
DB 237 SQFN-----PNNPO-QVYQPSQPGPWTTCPASNPLSHTSQQPNOGQHTSHVYMPI 287
QY 359 ADKTSKLMTEYKRRKQO-----HNTSAFPEYKVEKKGP-----LGR 398
DB 288 SSPTTSPPTTHSSGSSSOSSAHSOYNIONISTGPRKNOIEKLEPPORNNSSKLRSSSPR 347
QY 399 -----GODSEPLORPO-----HLMDOGQMRHSFSA-----GPELLRODKRPR 435
DB 348 TSSSTSVNSQTLNRNPQTYVIAASPRNTDELMRSQPKYVISAANAATGDEQVWRNPTL 407
QY 436 SGTSGSSLSVSRBAEQI-----QAMTNVLTVLNQIILPDQFT----- 477
DB 408 FISTNSASASARMSGQVSMGPAFIHHHPKRSRAIGNSATSPRVVYVTOENTRYTEKIT 467
QY 478 ---ALQPAVFPICISQLTCHVTDI 497
DB 468 VSPKPPAVSPGVSPTELTNL 490

RESULT 15

US-10-171-311-113
; Sequence 113; Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Gatti, Karen
; APPLICANT: Ganavayaru, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-113

Query Match 3.5%, Score 95.5; DB 9; Length 1713;
Best Local Similarity 19.4%, Pred. No. 44;
Matches 111; Conservative 84; Mismatches 207; Indels 169; Gaps 28;

QY 36 LPPDEKPNGRKTKSVSFREIYVS--LSHQVLLONLVDILLEEFYK-GPSPGEKTIQV 92
DB 1165 LPPTD-----HLQASFGFOTFQPSGILLDHQWTRMLQVTELDGYIELSTDSGGPIFKSP 1220
QY 93 EAKLAGFLRYISMQNLAVIFDILLDSYRTAREFPSPGLKCL-----KVSGIGCA 144
DB 1221 QTYMDGLLHYVS-----VISD-----NSGLRLIIDQLLRNKRKLHMISS 1261
QY 145 ANLYROSAMSF-----NIYFHALVCAVLTNOETTTAEQVK-----VLEF 184
DB 1262 ROSLRIGSNFEGCISNVFVQRL--SLSPVLDLTSNLSKRDVSLGCSLKKPPFLMLLK 1319
QY 185 DDERSTDSQOCSEDEDFEETAOVSPPRGKEKROW-ARMPLLSVQ-----P 232
DB 1320 G-----STRENKTKTRINOLLQDTPYASP-----RSYKWWODACSPLEKTOANHGAQFGDIP 1373
QY 233 VSNADWVWLKRLKLCMELCNNTYIOMHLENCMEEPPIF-KGDPEFILPSFQESSSTP 291
DB 1374 TSHL-----LEKLPQELKPRSQFAVDQMTSSRGLVHTGTKNSFMALYLSKGLV 1425
QY 292 STGFSGKETPSEDDRSQSR-----HMGSSLSKAGGDDLPPSPKVEKK 338
DB 1426 FALGTGDKKL-----RISKSKCNDGKWHTVFGDGK-----GRLVY----- 1464
QY 339 DPSRKKEWENAGNKITYMAADKTSKLMTEYKRRKQOHNLSAPPEKVEYKKG--EPLG 396
DB 1465 DGLRARE-GSLPGNSTISIRAPVYLGSPSPGKPKSLPNSVCGCLKNFOLDKPLYPSS 1523
QY 397 PRGODSPL--LQRPQHLMDG--QMRHSFAGPEL-LRODKRPR----- 435
DB 1524 SFGVSSCLGCPLEKGIYFSEEGHVLAHSLVLPREFLVSPRSRLTGILIHGSPG 1583
QY 436 -----SGTSGSSLSVSRBAEQIQAMTNVLTVLNQI--QILPDQT 475
DB 1584 KHLGVYLEAGKVTAASMDSGAGTSTSVTPKOSLCDQ--WHSVAVTIKQHLILLEDTPSS 1642
QY 476 FTALQPAVFPICISQLTCH-----VTDIRV 499
DB 1643 YTAGQIPPPASTQEPHLGAPANLTTLRI 1673

Search completed: June 17, 2003, 12:02:27
Job time : 58.3646 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:06 ; Search time 59.9151 Seconds
(Without alignments)
831.137 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRAMAQQVEMLDTCSPKTR.....VROAVREMLGRVGRVYDIIV 518

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	4.7	2278	1 S56274	FAB1 protein - yeast
2	124	4.6	2150	1 S27802	zinc finger protein
3	124	4.6	2150	2 T19450	hypothetical protein
4	113	4.2	482	2 T22981	hypothetical protein
5	111	4.1	781	2 T41551	hypothetical protein
6	111	4.1	1148	2 T09073	splicing factor S1
7	110	4.1	829	2 A33166	colorectal tumor s
8	109.5	4.1	587	2 T24103	hypothetical protein
9	108.5	4.0	2783	1 A41948	alpha-fetoprotein
10	108	4.0	506	2 T41788	global transactiva
11	108	4.0	1173	2 T25985	hypothetical protein
12	108	4.0	1912	2 T29088	vitellinogenin I pre
13	106.5	3.9	2617	2 AE2136	peptide synthetase
14	105	3.9	506	2 B72855	global transactiva
15	104.5	3.9	830	2 T16236	hypothetical protein
16	103.5	3.8	1304	2 T19397	hypothetical protein
17	103	3.8	453	2 T22153	hypothetical protein
18	103	3.8	1306	2 T13592	hypothetical protein
19	103	3.8	1446	2 T15212	hypothetical protein
20	103	3.8	1502	2 T42216	multidrug resistan
21	102.5	3.8	439	2 B84774	hypothetical protein
22	102	3.8	457	1 A28468	chromogranin A pre
23	101.5	3.8	621	1 S59632	endo-1,4-beta-xyla
24	101.5	3.8	682	1 B84021	two-component sens
25	101.5	3.8	2101	2 A42184	nuclear mitotic ap
26	101	3.7	859	2 T01461	hypothetical prote
27	101	3.7	1099	2 A59311	myosin VIII, ZMM3
28	101	3.7	1198	2 T20262	hypothetical prote
29	101	3.7	1203	2 C89217	protein C55A6.2 [1

30	100.5	3.7	453	2 S50248	SSF2 protein - yea
31	100.5	3.7	484	2 S60943	RND3 protein - yea
32	100.5	3.7	947	2 G70657	probable adl - Myc
33	100	3.7	1268	2 A44125	high density lipop
34	99.5	3.7	342	2 A56352	homeotic protein H
35	99.5	3.7	649	2 AG0710	DNA topoisomerase
36	99.5	3.7	1534	2 A56734	ribosome receptor,
37	99	3.7	456	2 S62962	hypothetical prote
38	99	3.7	822	2 T41622	probable ABC trans
39	99	3.7	1207	2 D84781	hypothetical prote
40	98.5	3.7	450	2 T21931	hypothetical prote
41	98.5	3.7	739	2 T21431	hypothetical prote
42	98.5	3.7	960	2 S44812	F4489.6 protein -
43	98.5	3.7	986	2 T41809	DNA polymerase orf
44	98.5	3.7	1750	2 E86151	hypothetical prote
45	98	3.6	692	2 E96841	hypothetical prote

ALIGNMENTS

RESULT 1

S56274
FAB1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YFR019w
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S56274; A55997
R:Murakami, Y.; Naitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sas
submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces
A:Reference number: S56186
A:Accession: S56274
A:Molecule type: DNA
A:Residues: 1-2278 <MOR>
A:Cross-references: EMBL:D50617; NID:9836685; PIDN:BAA09258.1; PID:9836774; GSPDB:GNO
R:Yamamoto, A.; Koshland, D.
submitted to GenBank, August 1993
A:Description: FAB1: a novel gene required for endocytic-vacuolar pathway and nuclear
A:Reference number: A55997
A:Accession: A55997
A:Molecule type: DNA
A:Residues: 1-2274, 'R', 2276-2278 <YAM>
A:Cross-references: GB:U01017; NID:9392993; PIDN:AAA81360.1; PID:9398498
C:Gene: SCD:FAB1; MIPS:YFR019w
A:Gene: SCD:FAB1; MIPS:YFR019w
A:Superfamily: FAB1 protein; 1-phosphatidylinositol-4-phosphate 5-kinase homology
F:1911-2266/Domain: 1-phosphatidylinositol-4-phosphate 5-kinase homology <PIK>

Query Match 4.7% Score 126 DB 1 Length 2278;

Best Local Similarity 21.5% Pred. No. 0.98; Matches 108; Conservative 77; Mismatches 148; Indels 170; Gaps 27;

OY	19	KTPNNF--DHAQC--OLITEL-----PDE--KNGCHKSVSPREIVS	58
DB	1330	KCPHDFKDHVKYGVNDLVRLKXSDLEHVELTPPKIKMKHIDIKLVE-----	1382
OY	59	LLSHQVL--LQNLVDLLE--EPVKSPGGEKTIQVPEAKLACFLAYISQNLAVFDL	114
DB	1383	LVYKILKINNFSGVLSRLERIKLDSMTKDV-----LSQAKIIEKSNATBEOK	1434
OY	115	L-----LDSYRTAREFDPSPGKCLLKVKVSGIGGANLYROSAMFNIFYFHLVCAVL--	167
DB	1435	LMLODLDFVADSPCDHLPINLVIK-----SLV-DKAVMNSTFAIFAKSVLPSE	1484
OY	168	TNQETITAEQYKVLFEDEDERSTDSQOCSEDEDFEETAAQVSPPKGEKRRQRRAMP	227
DB	1485	TDISIRITAKQKLKLFY-----DSRRKDSDDKSLHDEKAKTKRPEKNE-----LPL	1530
OY	228	LSVQPSNADWVWLKRLKLCMELCNNTYIOMHLDLENCE-----EP-----PIFK	275

Db 1531 EGLDVE-----KPKIDSKNTTENRDTNEPQNAVITTFKD 1567
Qy 276 DPFILPSFOSE-STSTPGSGFKETPSEDNRQ-SREHMGESLTKAGGDLPLPSP 333
Db 1568 DFLIPNPGSHLVTPQSSASSSLTPQTEERPISGCIQMT-----1613
Qy 334 KVEKKDPSRKKEWMENAGNKITYMAADKT---ISKLMTEYKKRKQ-----OHNLISA 381
Db 1614 ---HDKSTRP-----NIRKMSDSSLQGLASLANEYSKNNKVSKLATFFQMHFDA 1661
Qy 382 FPKVEYKKEGEP-----GPRGDS-PLLRPQHLMDQGMHRSFSGPELRLRDK 432
Db 1662 LSKFELEERERERLQLNKKDKQAIRLQSTPIVEIYKNVKD-----AVDE 1706
Qy 433 RPRSGTGSLS-SVSVRDPAEAOI 454
Db 1707 PLHSRSSGNLSANVKTLEAPV 1729

RESULT 2
zinc finger protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000
C:Accession: A45172: S27802
R:Klein, R.D.; Meyer, B.J.
Cell 72, 349-364, 1993
A:Title: Independent domains of the Sdc-3 protein control sex determination and dosage
A:Reference number: A45172; MUID:93161411; PMID:8431944
A:Accession: A45172
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2150 <KLE>
A:Cross-references: EMBL:M85149; NID:G156440; PIDN:AAA28144.1; PID:G156441
A:Experimental source: strain N2
A:Note: sequence extracted from NCBI backbone (NCBIF:124842)
C:Genetics:
A:Gene: sdc-3
A:Introns: 70/1; 102/1; 153/3; 258/1; 305/1; 1283/1; 1398/1; 1703/2; 1853/1; 1887/1; 190
C:Superfamily: zinc finger protein sdc-3
C:Keywords: DNA binding; nucleus; zinc finger

Query Match 4.6%; Score 124; DB 1; Length 2150;
Best Local Similarity 20.6%; Pred. No. 1.3; Mismatches 166; Indels 210; Gaps 24;
Matches 114; Conservative 63;

Qy 33 IEIPDEKPNHGHTKKSVSFREIVSL-----SHOVLQ---NLVDILLEEFVK 79
Db 1177 IERLSEDLGLDQSEDISFEIQLVDLLESGVEQVQSVTISRSTSFESLVEDPEE 1236
Qy 80 GP-----SPGEKTIQ-VPEAKLGLFLYISMQNLAVFDLLDSYRKAREFDSPGLKC 133
Db 1237 HPEDLPVSASEKANNQIVPEVEGVSVPVINO-----EENVVSEG--P 1279
Qy 134 LKRVSGIGGANLYRQSAFNFYFHALVCALVINO-----ETITAEQ-----177
Db 1280 TLQGGSSIPSSSHY-----TVDELIGTESPGREATETPAEESPKKSGKT 1326
Qy 178 -----VKYLFEDDERSTDSQOCSEDEDIFEETAQVSPRPGKEKQRRAMP--L 228
Db 1327 TRGRPKRYKENLKRIPRRCQKEBAHEPE-VVEEOQVPEVPEVHBPVAPAOQL 1385
Qy 229 SVQPVSNADWVWLKRLKLMELCNNTIQMHLDLENMEEP-----PIFGDPFFILPS 283
Db 1386 ETEPI-----EQDIEPDKVEFEPIEA-----LPL 1410
Qy 284 FQSSSTPSTGSGFKETPSED-----RSOSREHMG 315
Db 1411 FETSPVAPRPGNITSRAHSSDDVQVYISSETDPNGPILVQVONDKLTAYQSTELLG 1470
Qy 316 ESLSLKAGG-----DLLPSPKVEKKDPSRKKEWMEN--AGNKITYMAA 359
Db 1471 EYGLDEAGAPSPSEIIVHDEVLODEVLPKSSKKRGRKKTTPPHIAARVFT---1527

Qy 360 DKTISKLMTE-----YKRRQOHNLSA-----FPKEVVEKKGEF-----LGR 398
Db 1528 --SISK--TEIEELAPPTQOSRKRMANVSSSEATATRRQKRAVEEENDSDVSRVLTP 1583
Qy 399 GQDSFLLORPOLMDQGMHRSFSGPELRLQDKRPNRSGTSSLSVS-----446
Db 1584 PEDLHETERPCHVGEEGFETPSLRTGRE-----STASSVKTSSRSKRLFLSKNMP 1632
Qy 447 VRDAEOIQAWTN 459
Db 1633 VPRMRIOQAQTN 1645

RESULT 3
T19450
hypoetical protein C25D7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C:Accession: T19450
R:Ainscough, R.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19125
A:Accession: T19450
A:Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-2150 <WTL>
A:Cross-references: EMBL:Z81039; PIDN:CA802774.1; GSPDB:GNO0023; CESP:C25D7.3
A:Experimental source: clone C25D7
C:Genetics:
A:Gene: CESP:C25D7.3
A:Map position: 5
A:Introns: 70/1; 102/1; 153/3; 258/1; 305/1; 1283/1; 1398/1; 1703/2; 1853/1; 1887/1;
C:Superfamily: zinc finger protein sdc-3

Query Match 4.6%; Score 124; DB 2; Length 2150;
Best Local Similarity 20.6%; Pred. No. 1.3; Mismatches 166; Indels 210; Gaps 24;
Matches 114; Conservative 63;

Qy 33 IEIPDEKPNHGHTKKSVSFREIVSL-----SHOVLQ---NLVDILLEEFVK 79
Db 1177 IERLSEDLGLDQSEDISFEIQLVDLLESGVEQVQSVTISRSTSFESLVEDPEE 1236
Qy 80 GP-----SPGEKTIQ-VPEAKLGLFLYISMQNLAVFDLLDSYRKAREFDSPGLKC 133
Db 1237 HPEDLPVSASEKANNQIVPEVEGVSVPVINO-----EENVVSEG--P 1279
Qy 134 LKRVSGIGGANLYRQSAFNFYFHALVCALVINO-----ETITAEQ-----177
Db 1280 TLQGGSSIPSSSHY-----TVDELIGTESPGREATETPAEESPKKSGKT 1326
Qy 178 -----VKYLFEDDERSTDSQOCSEDEDIFEETAQVSPRPGKEKQRRAMP--L 228
Db 1327 TRGRPKRYKENLKRIPRRCQKEBAHEPE-VVEEOQVPEVPEVHBPVAPAOQL 1385
Qy 229 SVQPVSNADWVWLKRLKLMELCNNTIQMHLDLENMEEP-----PIFGDPFFILPS 283
Db 1386 ETEPI-----EQDIEPDKVEFEPIEA-----LPL 1410
Qy 284 FQSSSTPSTGSGFKETPSED-----RSOSREHMG 315
Db 1411 FETSPVAPRPGNITSRAHSSDDVQVYISSETDPNGPILVQVONDKLTAYQSTELLG 1470
Qy 316 ESLSLKAGG-----DLLPSPKVEKKDPSRKKEWMEN--AGNKITYMAA 359
Db 1471 EYGLDEAGAPSPSEIIVHDEVLODEVLPKSSKKRGRKKTTPPHIAARVFT---1527
Qy 360 DKTISKLMTE-----YKRRQOHNLSA-----FPKEVVEKKGEF-----LGR 398
Db 1528 --SISK--TEIEELAPPTQOSRKRMANVSSSEATATRRQKRAVEEENDSDVSRVLTP 1583
Qy 399 GQDSFLLORPOLMDQGMHRSFSGPELRLQDKRPNRSGTSSLSVS-----446


```

OY 214 RGEKRWKRAMPLLSVQVPSNADWVWLVKRLHKLMCLNNYIOMHLDLENCEMEPIIF 273
DB 417 -----EVEPSVNA-----LKQNMENSVTHCS-----ENNM--PSSD 346
OY 274 KGDFEFLPSFOSES-----STPSTGFGSGKETPSEDROSOREMG 315
DB 447 LADEKEVEVQSPSEPKDTIDTKKPRTRRSRFRSHSPSTTWSBNKDTPEKKRPQS----- 501
OY 316 ESLSLKAGGDLLEPPSPKVEKKDPSRK-----KEWENAGNKITYMAADKTSKLMTE 369
DB 502 -----PSPRRETKGKSRKSSQSPSPKESARGRKRSQSPKKDIA----- 541
OY 370 YKKRQOHNLSAFPKEVVEKGEPLGRGQDSPLLQRPQILMOGQKRHSFSAGPELLR 429
DB 542 -RERQOSRSPKRDPTRESRSESLSPRRETS-----RENKRSQPRVKDS--SPGKRS 594
OY 430 QDKRPSGSGTSSLSVSVDAEAQIOWT 458
DB 595 QSRERESDRDQR-----RERERTRRKWS 618

```

RESULT 7

A33166

colorectal tumor suppressor protein - human

C:Species: Homo sapiens (man)

C:Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 21-Jul-2000

C:Accession: A38434; A33166

R:Kinzler, K.W.; Nishitoh, M.C.; Vogelstein, B.; Bryan, T.M.; Levy, D.B.; Smith, K.J.; Pr

Science 251, 1366-1369, 1991

A:Title: Identification of a gene located at chromosome 5q21 that is mutated in colorect

A:Reference number: A38434; MUID:91164855; PMID:1848370

A:Accession: A38434

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-829 <KIN>

A:Cross-references: GB:M62397; NID:g181034; PIDN:AA52069.1; PID:g181035

C:Genetics:

A:Gene: GDB:MCC

A:Cross-references: GDB:128163; OMIM:159350

A:Map position: 5q21-5q21

```

Query Match 4.1%; Score 110; DB 2; Length 829;
Best Local Similarity 18.4%; Pred. No.3.8;
Matches 92; Conservative 81; Mismatches 199; Indels 130; Gaps 19;

OY 81 PSEBEKTI-----OPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFTSPGICKLLK 136
DB 307 PSTGELSTSSSSNDIPIKIAERYK-----LSKTRSESSSSDRPVLSGS 349
OY 137 KVSIGGANLYRQSAWSE-----NIYFHALVCAVLINQETTTAQVKVLEED 185
DB 350 EISSIGVSSVAEHLAHLSDCSNIQETFTLYSHG-----SAISESKIRE--FEV 398
OY 186 DERSTSSQCCSEDEDFEETAGVSPRGKEKROWRAMPLLSVQVPSNADWVWLVRLL 245
DB 399 ETEFLNSIHEHLKQONDLITLLE-----ECKSNARMGLVKYSNMTALRLAQY 451
OY 246 HKLQELCNNTIOMHLDLENCEMEPIIFKGDPEFLPSFOSESSTPSTGSGKETPSE- 304
DB 452 SEQCEIAVEYELL-----ALAESEQSLILQGFRAAGVSSPGDQSDENITQM 498
OY 305 -----DORSQREHMGESLSLKAGG--GDLLLPSPKVEKKRPSRKKEWENAGNKITYMA 358
DB 499 LKRAHDCRKTIAENAKALMLKLDSSCGAFVAVAGCSVQ-----WESLSSNSHST 549
OY 359 ADKTIKSLMTEYKRRKQOHNLSAFPKEVVEKKG-----EPLG-----PRGOD 401
DB 550 TSSIASQDTEFTEDEQ-RLKDYIQQLKNDRAAVKLTMLESTHIDPLSYDVKPRKDS 608
OY 402 SPL-LQRP---QHLMDQGMHRHSFAGPELLRQDKRPPSGSGTSSLSVSVDAEAQIOWT 457
DB 609 QRLDLNNAVLMQELMAMKEWAEELKAQLYLLEKEKK-----ALTEKLSTRQEQAY 660

```

```

OY 458 TNNVLTVLNOIQLPDQFTALQ-----PAVEPCIS-----QLTCHVDIR 498
DB 661 LVHIEHLKSEVEGEQRMRSLSSTSSGSKDKPGKCADASPALSIAELRTCSENELA 720
OY 499 VR--QAVREMLGRVRYDII 517
DB 721 AEFTHAIRREKKRLKARVOELV 741

```

RESULT 8

T24103

hypothetical protein R102.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24103

R:Berkas, M.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19841

A:Accession: T24103

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-587 <MIL>

A:Cross-references: EMBL:Z70309; PIDN:CAA94361.1; GSPDB:GN00022; CESP:R102.7

A:Experimental source: clone R102

A:Map position: 4

A:Introns: 19/2; 111/3; 147/1; 270/2; 407/1; 481/3; 529/2

```

Query Match 4.1%; Score 109.5; DB 2; Length 587;
Best Local Similarity 20.7%; Pred. No.2.5;
Matches 84; Conservative 60; Mismatches 132; Indels 129; Gaps 19;

OY 77 FVKG---PSPBEKTIQVPEAKLAGFLRYISMQNLAVIFDILLDSYRTA--REEDTSPGL 131
DB 6 FRKGQVLPISSKRRTLOMLEKKAFAVRSILVREVG-----FLDRHDTCHEREVQLTPEDD 60
OY 132 KCLLKVSGIGGANLYRQSAWSENIYFHALVCAVLINQETTTAEQVKVLEPDERSTD 191
DB 61 QKLIKLVK--EGETNMLEKGSKLFDLL--TMSGSLI-----VRGATIEFNPNNE 107
OY 192 SSQCCSDEDEDFEETAGVSPRG--GKEKROWRAMPLLSVQVPSNADWVWLVRLLH 246
DB 108 TMDQDNEEAIVIEDLVOLKSKKEKEKKKTKTERMDALDELTHVTSNDRIISLSR-- 165
OY 247 KLCMELCNNTIOMHLDLENCEMEPIIFKGDPEFLPSFOSES-----TPSTGFSGKET 301
DB 166 -----EDDSQREHMGESLSLKAGGDLLEPPSPKVEKKRPSRKKEWENAGNKITYT 189
OY 302 PS-----EDDSQREHMGESLSLKAGGDLLEPPSPKVEKKRPSRKKEWENAGNKITYT 356
DB 190 VSVQPTLDDMDTQK-----YVKKRSPRHK--PSTRE-----HT 224
OY 357 MAADKTIKLMTE--YKKRQOHNLSAFPKEVVEKKG-----GEPLG 396
DB 225 LSVQKTVEDMDTKQKVVKKPSSRPKRPSTHEVLAKKKSAATISKETCOLHTKFGKLFG 284
OY 397 --PRGQDSP-----LQRPQHLMDQGMHRHSFAGPELLRQDKR 433
DB 285 TVTEPNNSPYQNLIDLSPRPSKYQDNQPMFLPSPINPTATNEIKK 329

```

RESULT 9

A11948

alpha-fetoprotein enhancer-binding protein - human

N:Alternate names: ATFBI protein

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999

C:Accession: A11948

R:Moringa, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaki, T.

Mol. Cell. Biol. 11, 6041-6049, 1991

A:Title: A human alpha-fetoprotein enhancer-binding protein, ATFBI, contains four hom

QY	219	ROMAHRMLP-----LSQVPSNADWMLVKRLKHCIMELCNMYIQMHLDLENMCMEPPIFGK	275
Db	247	-SPAKGPIILNKKPARISSSSNPKFYOLMNR-----TIAPVSTSS	284
QY	276	DPFLLPFSQSSSTPTSTGFGSGKTEPSEDDRRSQSHHGESISLAKAGDDLPLP----	331
Db	285	RP--PIPIFPVVTASRP-----RPND-----RPLRPMAPKLPVK-----VFIPSLARK	324
QY	332	-----SPREVEKKDPSRK--KEWMENAGNKITYMAADKTSIKLMTYEKKRRQOHNLSA	381
Db	325	INSSAVPLRDRKRIELKDKPAELKD--SNIVLKGITIKRPFVIEIETPRPPFKKQOEQ----	379
QY	382	FPREKVEKKGPIG--PGQDSPLLP-----	408
Db	380	--EHTNGEKKKEIILDEPIIDEPRTMRPRHEHCAIFNNLSIPKPRHGENTTSIRKVTIMPK	457
QY	409	-QHLMDOGOMHRSFAGPELLRQDKRPRS	436
Db	438	SMHLNKGE-----SAGPSLFATNRPRS	461

RESULT 12

vitellogenin I precursor [validated] - chicken
N:Contants: lipovitellin I; lipovitellin II; phosvitin; yolk glycoprotein 42K
C:Species: Gallus gallus (chicken)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: J29088; S53681
R:Mdbuchl, N.; Yamamura, J.; Adachi, T.; Aoki, N.; Nakamura, R.; Matsuda, T.
submitted to the EMBL Data Library, November 1996
A:Description: cDNA cloning and estrogen-induced expression of chicken vitellogenin I.
A:Reference number: 220557
A:Accession: J29088
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1912 <MAB>
A:Cross-references: EMBL:D89547; PIDN:BA13973.1
A:Experimental source: liver
R:Yamamura, J.; Adachi, T.; Aoki, N.; Nakajima, H.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1244, 384-394, 1995
A:Title: Precursor-product relationship between chicken vitellogenin and the yolk prote

Query Match	4.0%;	Score 108;	DB 2;	Length 1912;
Best Local Similarity	20.3%;	Pred. No. 18;		
Matches 73; Conservative	39;	Mismatches 137;	Indels 110;	Gaps 10;

```
QY      176   QVAVKVLFEEDREDSGQQSSDEDEDI FEETAQVSPPRGKEKRQWRARMLPLLSQVPYN 235
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1076  KRKKILLDDTDNQATRNRSRSSSSASSISESESTTSPSSSDSNRAS----- 1122
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      236   ADWAWLVKRLHKILCMELCNNTIOHLDLENCMEEPPIRGDPFFILPFSOF- 287
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1125  -----GGDPOILNKRKSANCKRFP 1144
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      288  -----SPTPTGTGTSCKTTBPEDDRSGSRHEMGBSLSLKAGGDLLPPS 332
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1147  FGDSSSSGSSSSSSSSSSSSSSSSSKSSSSSDSSSSSSSSSSSSSSSSSSSSSKSS--RS 1202
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      333   PKVEKIDPSRKKEWMENAGNKITYMMAAKTISLKTETVKRRROHNLSAPREXVVEKKG 392
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1203  SKSMRSSSSSNSKROSSSSSSSKNSKNGSSSSSSSKAGCTQKAKKQKCTTTFPHASAAE--G 1266
          |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      393   EPLPGCGDPSFLLRPOHLMDQGWRHSFAGPELLRDOKRPBGSTGS--SLESVARD 449
```

Db 1261 E-----RSVH- -EQKPEIQSSSSSSSSASSNSKSTSTSSSSSESSSVHRQ 1305

Qy 450 AEAQIQAIVNVLTVLNOIQI- -LPDQTFALQPAVFPCISQLTCHVYDIFVRQAVREM 506

Db 1306 WKQREAEFTKRVKQSFNSHSSYDIPNEHETVL- -EKVY-----RLFRSRKHTHW 1352

RESULT 13

peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence-revision 14-Dec-2001 #text-change 30-Jun-2002
C:Accession: AE2136
R:Kaneaki, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaamoto, S.; Watanabe, A.; Iriy
Nakanaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Residues: 1..2617 <R0>
A:Cross-references: GB:BA0000019: PID:gl713137: GSPDB:GN00179
A:Experimental source: strain PCC 7120
A:Genetics:
A:Gene: all2644
A:Superfamily: peptidase synthetase ppsd: acetate-CoA ligase homology: acyl carrier protein
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:1025/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
F:1068/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match	3.9%;	Score 106.5;	DB 2;	Length 2617;
Best Local Similarity	20.2%;	Pred. No. 36;		
Matches 75; Conservative	49;	Mismatches 125;	Indels 123;	Gaps 16;

[illegible]

RESULT 14

global transactivator-like protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: B72855

R:Avres, M.D., Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus
A:Reference number: A72850; MUID:94303173; PMID:8030224
A:Accession: B72855
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-506 <AVR>
A:Cross-references: GB:L22858; NID:g510708; PIDN:AA66672.1; PID:g559111
C:Genetics:
A:Gene: Ac-GTA

	Query Match	3.9%	Score 105;	DB 2;	Length 506;	
	Local Similarity	22.3%	Pred. No. 4.5;			
	Matches	69;	Conservative	49;	Mismatches	91;
					Indels	100;
					Gaps	18
OY	126 DTSPGLKELKRVSGISGANILYR-QSASMFNIFHAUVCALVTNGETITTAQVKKVFEE	184				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	222 DSTNRKIISKIKIV-----LKRRKSSEISSNIPIKH-----TVEYV-HVNPN	260				
OY	185 DDERSTDSQQCCSEDEDIFEETAOVSPPRCKEKQMARRPRLSVOPYSNADWYLVKR	244				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	261 EEETIYLVDKLCESEE-----AVYKAVAARENNALSIR-LQOMOHV-----LWLIIK	306				
OY	245 LHKICMELCNVYIQHLDLENOMEERPIFEKGDPF-----FILPS	283				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	307 LRQIC--CHRYLANH-GKNILETNDCFKMDYSSCKCRVLVDLYDLILNTSNDKIILVS	361				
OY	284 -----FGS-EESTPSTGFSGSK-----ETPSEDRSQSREHGESLSLKAGG	324				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	362 QMWELYLKIFENFFKKNKNIATLTMTQLAVEEDRIIALETTFND--AANTQIRILLISIKCGG	419				
OY	325 GDL-----LLRP-SPRYEKKDPSRKKEWMENAGNIKYTM--AADRTISKIMTEYK	371				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	420 VGLNLIGGNHIYMLERPWNPOIELOADRISMGCOTKNTYYVKMLNVEDNSIEK----	475				
OY	372 KKKQOHNLS	380				
	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
Dd	476 KORODKKTA	484				

RESULT 15
T16236
hypothetical protein F32A6.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16236
R:Wu, X.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F32A6.
A:Reference number: Z18483
A:Accession: T16236
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <WU>
A:Cross-references: EMBL:U04009; NID:g1065447; PID:g1065450; PIDR:AAA81388.1; CESP:F32A6
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F32A6.3
A:Introns: 31/2; 69/3; 111/3; 146/2; 247/3; 290/2; 352/3; 416/3; 632/3; 666/1; 736/1; 80

Query Match	3.9%	Score 104.5;	DB 2;	Length 830;
Best Local Similarity	18.3%;	Pred. No. 9.9;		
Matches	66;	Conservative	64;	Mismatches 138;
			Indels	93;
			Gaps	15

```

OY      TRRAAAOQVEMLDTOSPK---TPNEDHOSQILTELPDEKPNNGHTKSVSRIVV 57
Db      379 RARRAASRLPLIGCEKREEMAVNPFEEVKCLTLAEVYDPGP---TIDPECYQAVLI 4353
OY      58 SILSHOY-----LLONLTY-DILLEEV-----KQSPGEKRTIQQYPEAKLAG 98
Db      436 AGLFNFKQFKRLTYQWSPDLYMTSTIIDFTQRIQOISKNSGLNADDEREYVLMODLAAH 4951

```

[illegible]

Search completed: June 17, 2003, 12:01:08
Job time : 63.1151 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Coriagen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:06 ; Search time 4.74231 Seconds
(without alignments)
831.137 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EIDERSTDSQQCSSEDEDI.....ETAGVSPPGKEKRWRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	29.6	707	2 A35804	nucleolin - human
2	64	29.6	712	2 JH0148	nucleolin - rat
3	63	29.2	713	2 A27441	nucleolin - Chinese
4	61	28.0	2359	2 A9616	unknown protein FI
5	60.5	28.0	512	1 WMBEXA	UL54 protein - hum
6	60	27.8	395	2 S66940	SGT1 protein - yea
7	59.5	27.5	206	2 S70004	hypothetical prote
8	59.5	27.5	767	2 S63182	hypothetical prote
9	59	27.3	586	2 C83262	hypothetical prote
10	59	27.3	707	1 DNMS	nucleolin - mouse
11	58.5	27.1	182	2 T46396	hypothetical prote
12	58	26.9	511	1 A48560	UL54 protein - hum
13	58	26.9	512	1 WMBEXA	UL54 protein - hum
14	58	26.9	796	1 A32434	abaa protein - Eme
15	58	26.9	1738	2 S20614	conserved hypotet
16	58	26.9	3839	2 T49799	related to TOM pr
17	57.5	26.6	88	2 D95264	hypothetical prote
18	57	26.4	340	2 T51386	probable protein w
19	57	26.4	396	1 A58938	surface protein w
20	57	26.4	2643	2 T29149	hypothetical prote
21	56.5	26.2	641	2 T17278	hypothetical prote
22	56	25.9	317	2 S41081	monomine transpor
23	56	25.9	1695	2 T19823	hypothetical prote
24	56	25.9	1711	1 A47392	chromodomain-helic
25	55.5	25.7	315	2 T15460	hypothetical prote
26	55.5	25.7	351	2 T32093	hypothetical prote
27	55.5	25.7	486	2 A89827	elastin binding pr
28	55.5	25.7	1465	2 S45628	DNA-directed DNA p
29	55.5	25.7	2135	2 T14602	variant-specific s

30	55	25.5	411	2 S48647	peptidylprolyl iso
31	55	25.5	426	2 T05676	hypothetical prote
32	55	25.5	731	2 T08855	nephrocystin - hum
33	55	25.5	765	2 E96558	hypothetical prote
34	55	25.5	1040	2 S50617	BEB1 protein - yea
35	55	25.5	1241	2 JU0466	potassium transpor
36	54.5	25.2	440	2 A42136	transcription fact
37	54.5	25.2	37	2 S49634	hypothetical prote
38	54.5	25.2	1070	2 T31069	collid-BMP-1 like
39	54	25.0	206	2 S43445	translational elonga
40	54	25.0	314	2 F84865	probable C2H2-type
41	54	25.0	328	2 T01225	hypothetical prote
42	54	25.0	328	2 B84545	hypothetical prote
43	54	25.0	332	2 C40646	endospore developm
44	54	25.0	44	2 A45988	dent matrix acid
45	54	25.0	529	2 S4669	f42h10.7 protein -

ALIGNMENTS

RESULT 1

A35804
nucleolin - human
N:Alternate names: phosphoprotein p100; protein B50; protein C23
C:Species: Homo sapiens (man)
C>Date: 23-Oct-1990 #sequence,revision 23-Oct-1990 #text_change 23-Jul-1999
C:Accession: A35804; S04631; A48138; A55996
R:Stivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
J. Biol. Chem. 265, 14922-14931, 1990
A:Title: Genomic organization and chromosomal localization of the human nucleolin gene
A:Reference number: A35804; MUID:90368666; PMID:2394707
A:Accession: A35804
A:Molecule type: DNA
A:Residues: 1-707 <SR1>
A:Cross-references: GB:W60858; GB:J05584; NID:q189305; PIDN:AAAS9954.1; PID:q189306
R:Stivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
FEBS Lett. 250, 99-105, 1989
A:Title: Cloning and sequencing of the human nucleolin cDNA.
A:Reference number: S04631; MUID:89290043; PMID:2737305
A:Accession: S04631
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-707 <SR2>
R:Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cecch, T.R.
Mol. Cell. Biol. 13, 4301-4310, 1993
A:Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUG/G/C) an
A:Reference number: A48138; MUID:93309464; PMID:8321232
A:Accession: A48138
A:Molecule type: protein
A:Residues: 458-474 <ISH>
A:Experimental source: Hela cell nuclei
A>Note: Sequence extracted from NCBI Backbone (NCBI:P134645)
R:Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.
Biochemistry 33, 14696-14706, 1994
A:Title: Major cell surface-located protein substrates of an ecto-protein kinase are
A:Reference number: A55996; MUID:95086063; PMID:7993898
A:Accession: A55996
A:Molecule type: protein
A:Residues: 231-236;349-362;399-403;458-461;655-656; 'X', 658-660 <JOR>
A:Experimental source: surface-labelled Hela cells
C:Genetics:
A:Gene: GDB:NCL
A:Cross-references: GDB:125908; OMIM:164035
A:Map position: 2q12-2qter
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding
F:308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F:394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F:487-550/Domain: ribonucleoprotein repeat homology <RRM3>
F:573-634/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 29.6%; Score 64; DB 2; Length 707;

R:McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.
 J. Gen. Virol. 72, 3057-3075, 1991
 A:Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
 A:Reference number: J01494; MUID:92113549; PMID:1662697
 A:Accession: J01498
 A:Molecule type: DNA
 A:Residues: 1-512 <MCG>
 A:Cross-references: GB:D10471; DDBJ:D01128; NID:g221784; PIDN:BA01269.1; PID:g221788
 C:Genetics:
 A:Gene: UL54
 C:Superfamily: varicella-zoster virus gene 4 protein
 C:Keywords: transcription regulation

Query Match 28.0%; Score 60.5; DB 1; Length 512;
 Best Local Similarity 34.1%; Pred. No. 10;
 Matches 14; Conservative 7; Mismatches 13; Indels 7; Gaps 1;
 OY 3 DERSTSSQCCSEDEDFE-----ETAAVSPPRGKERK 36
 Db 34 DDPSSSSGECSSDEDEDFEPCGGGAIDAIAPKGP 74

RESULT 6
 S66940
 SGT1 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein 02805; protein YOR057W
 C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 05-Nov-1999
 A:Accession: S66940
 R:Bonin, C.; Bolotin-Fukuhara, M.; Dajman-Fornier, B.; Dang, D.V.; Valens, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66929
 A:Molecule type: DNA
 A:Residues: 1-395 <BOH>
 A:Cross-references: EMBL:Z74965; NID:q1420194; PID:e251976; PID:q1420195; GSPDB:GN00015;
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:SGT1; GSSI: MIPS:YOR057W
 A:Cross-references: SGD:S0005583; MIPS:YOR057W
 A:Map position: 15R
 C:Function:
 A:Description: suppressor of G2 allele of SKP1

Query Match 27.8%; Score 60; DB 2; Length 395;
 Best Local Similarity 40.0%; Pred. No. 9;
 Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 OY 14 SSEDEDFEETAAVSPPRGKERK 38
 Db 371 STDWEDVSKGTVTSPEGMPEPKM 395

RESULT 7
 S70004
 hypothetical protein 206 - radish mitochondrion
 C:Species: mitochondrion Raphanus sativus (radish)
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 25-Apr-1997
 A:Accession: S70004
 R:Rankin, C.T.; Cutright, M.T.; Makaroff, C.A.
 Curr. Genet. 29, 564-571, 1996
 A:Title: Characterization of the radish mitochondrial nad3/rps12 locus: analysis of rec
 A:Reference number: S70000; MUID:96269916; PMID:8662196
 A:Accession: S70004
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-206 <RAN>
 A:Cross-references: GB:U43507
 C:Genetics:
 A:Note: the authors translated the codon GAT for residue 147 as Asn
 A:Genome: mitochondrion
 C:Keywords: mitochondrion

Query Match 27.5%; Score 59.5; DB 2; Length 206;
 Best Local Similarity 39.4%; Pred. No. 5.2;
 Matches 13; Conservative 5; Mismatches 12; Indels 3; Gaps 1;
 OY 8 DSSQCCSEDE---DIFETAAVSPPRGKERK 37
 Db 138 ESFHCEDDFELDVLEESFSDSPAPSPRK 170

RESULT 8
 S63182
 hypothetical protein YNL224C - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein NL269
 C:Species: Saccharomyces cerevisiae
 C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 19-Apr-2002
 A:Accession: S63182; S63190; S67374; S72092
 R:Duisterheft, A.; Floeth, M.; Fritz, C.; Heuss-Nelzel, D.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S62944
 A:Accession: S63182
 A:Molecule type: DNA
 A:Residues: 1-767 <DUE>
 A:Cross-references: EMBL:Z71500; NID:q1302244; PID:e239703; PID:q1302245; MIPS:YNL224
 A:Experimental source: strain S288C
 R:Pandolfo, D.; De Antoni, A.; Lanfranchi, G.; Valle, G.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63188
 A:Accession: S63190
 A:Molecule type: DNA
 A:Residues: 1-767 <PAN>
 A:Cross-references: EMBL:Z71500; NID:q1302244; PID:e239703; MIPS:YNL224
 A:Experimental source: strain S288C
 R:Pandolfo, D.; de Antoni, A.; Lanfranchi, G.; Valle, G.
 submitted to the EMBL Data Library, February 1996
 A:Description: DNA sequence of cosmid 14-5 from chromosome XIV.
 A:Reference number: S67355
 A:Accession: S67374
 A:Molecule type: DNA
 A:Residues: 1-767 <PAW>
 A:Cross-references: EMBL:Z69381; NID:q1183970; PID:e221958; PID:q1183990
 R:Pandolfo, D.; de Antoni, A.; Lanfranchi, G.; Valle, G.
 Yeast 12, 1071-1076, 1996
 A:Title: The DNA sequence of cosmid 14-5 from chromosome XIV reveals 21 open reading
 A:Reference number: S72073; MUID:97051596; PMID:8896273
 A:Accession: S72092
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-767 <PAF>
 A:Cross-references: EMBL:Z69381; NID:q1183970; PIDN:CAA93374.1; PID:q1183990
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
 C:Genetics:
 A:Cross-references: SGD:S0005168
 A:Map position: 14L
 A:Note: YNL224C

Query Match 27.5%; Score 59.5; DB 2; Length 767;
 Best Local Similarity 44.4%; Pred. No. 21;
 Matches 20; Conservative 4; Mismatches 12; Indels 9; Gaps 3;
 OY 1 EDDER---SFD---SSQCCSEDEDFE---ETAAVSPPRGKERK 36
 Db 181 EDDDSQNSPSTHDSISNESKVEDGDFEVDAAQSPDLRIK 225

RESULT 9
 C83262
 hypothetical protein PA3074 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 A:Accession: C83262
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
 Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 ; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: C83262
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-586 <STO>
 A:Cross-references: GB:AE004731; GB:AE004091; NID:9949171; PIDN:AMG06462.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3074

Query Match 27.3%; Score 59; DB 2; Length 586;
 Best Local Similarity 33.3%; Pred. No. 18;
 Matches 11; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

4 ERSTDSQCCSEDEDIFEEFAQVSPPRGKEKR 36
 Db 458 QROEQAAQQAQNGEKORERASQSPSSQR 490

RESULT 10
 DNMS
 nucleolin - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
 C:Accession: A29958; A40769; A56240; I84688
 R:Bourbon, H.M.; Lapeyre, B.; Amautic, F.
 J. Mol. Biol. 200, 627-638, 1988
 A:Title: Structure of the mouse nucleolin gene. The complete sequence reveals that each
 A:Reference number: A29958; MUID:88316930; PMID:3137346
 A:Accession: A29958
 A:Molecule type: DNA
 A:Residues: 1-707 <BOU>
 A:Cross-references: GB:X07699; NID:953453; PIDN:CAA30538.1; PID:953454
 R:Pasternack, M.S.; Bleier, K.J.; McInerney, T.N.
 J. Biol. Chem. 266, 14703-14708, 1991
 A:Title: Granzyme A binding to target cell proteins. Granzyme A binds to and cleaves nucleolin
 A:Reference number: A40769; MUID:91317840; PMID:1860869
 A:Accession: A40769
 A:Molecule type: protein
 A:Residues: 2-20, 'X', 22-24 <PAS>
 R:Yang, T.H.; Tsai, W.H.; Lee, Y.M.; Lei, H.Y.; Lai, M.Y.; Chen, D.S.; Yeh, N.H.; Lee, S.
 Mol. Cell. Biol. 14, 6068-6074, 1994
 A:Title: Purification and characterization of nucleolin and its identification as a transmembrane protein
 A:Reference number: A56240; MUID:94344117; PMID:8065340
 A:Accession: A56240
 A:Molecule type: protein
 A:Residues: 2-19; 558-567 <YAN>
 R:Bourbon, H.
 Gene 68, 73-84, 1988
 A:Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of
 A:Reference number: I48118; MUID:89121496; PMID:2906027
 A:Accession: I84688
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-44 <RES>
 A:Cross-references: GB:M2089; NID:g20011; PIDN:AAA3841.1; PID:9554246
 C:Comment: This housekeeping protein is involved in the synthesis, packaging, and mature
 C:Genetics:
 A:Insertions: 6/3; 45/3; 211/1; 273/1; 302/1; 349/2; 391/1; 432/3; 484/1; 525/2; 566/1; 608/1
 C:Superfamily: nucleolin; ribonucleoprotein repeat homology
 C:Keywords: DNA binding; duplication; nucleus; phosphoprotein; RNA binding; transcript
 F:310-375/Domain: ribonucleoprotein repeat homology <RRM1>
 F:311-316/Region: RNA-binding RNP2 motif
 F:349-356/Region: RNA-binding RNP1 motif
 F:396-458/Domain: ribonucleoprotein repeat homology <RRM2>
 F:397-402/Region: RNA-binding RNP2 motif
 F:431-438/Region: RNA-binding RNP1 motif
 F:468-551/Domain: ribonucleoprotein repeat homology <RRM3>
 F:489-494/Region: RNA-binding RNP2 motif
 F:524-531/Region: RNA-binding RNP1 motif
 F:570-634/Domain: ribonucleoprotein repeat homology <RRM4>

F:571-576/Region: RNA-binding RNP2 motif
 F:607-614/Region: RNA-binding RNP1 motif

Query Match 27.3%; Score 59; DB 1; Length 707;
 Best Local Similarity 32.4%; Pred. No. 22;
 Matches 11; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

1 EDDERSTDSQCCSEDEDIFEEFAQVSPPRGKE 34
 Db 193 EDEDEDEDEDEDEDEDEDEDEEVEVETTKAKKK 226

RESULT 11
 T46396
 hypothetical protein DKFZp434i1820.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46396
 R:Platenhaider, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223031
 A:Accession: T46396
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-182 <AAA>
 A:Cross-references: EMBL:AL137730
 A:Experimental source: adult testis; clone DKFZp434i1820
 C:Genetics:
 A:Note: DKFZp434i1820.1

Query Match 27.1%; Score 58.5; DB 2; Length 182;
 Best Local Similarity 34.8%; Pred. No. 6.1;
 Matches 16; Conservative 8; Mismatches 17; Indels 5; Gaps 1;

1 EDDERSTDSQCCSEDEDIFEEFAQVSPPRGKEKRR 41
 Db 11 EDELLSSESESESTDEDDORANKMLMELANLQPKRKTKQRHVR 56

RESULT 12
 A48560
 UL54 protein - human herpesvirus 1 (strain HPEM)
 C:Species: human herpesvirus 1
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Feb-1997
 C:Accession: A48560
 R:Rosen-Wolff, A.; Frank, S.; Raab, K.; Moyal, M.; Becker, Y.; Darai, G.
 Virus Res. 25, 189-199, 1992
 A:Title: Determination of the coding capacity of the BamHI DNA fragment B of *apathogen*
 A:Reference number: A48560; MUID:93070559; PMID:1332274
 A:Accession: A48560
 A:Molecule type: DNA
 A:Residues: 1-511 <ROS>
 A:Note: sequence extracted from NCBI backbone (NCBIN:117573, NCBI:117574)
 C:Genetics:
 A:Gene: UL54
 C:Superfamily: varicella-zoster virus gene 4 protein
 C:Keywords: transcription regulation

Query Match 26.9%; Score 58; DB 1; Length 511;
 Best Local Similarity 40.7%; Pred. No. 21;
 Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

3 DERSTDSQCCSEDEDIFEEFAQVSP 29
 Db 33 DDLESDSSGECSSDEDEMDPHGEGDP 59

RESULT 13
 WMBEY4
 UL54 protein - human herpesvirus 1 (strain 17)
 C:Species: human herpesvirus 1
 C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
 C:Accession: I30089

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:41 ; Search time 2.4612 Seconds

(Without alignments)
690.935 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDDERSTDSQCCSEDEDI.....EETAGVSPRGKEKROWRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	64	29.6	706 1 NUCL_HUMAN	P19338 homo sapien
2	64	29.6	712 1 NUCL_RAT	P13383 ratius norv
3	63	29.2	713 1 NUCL_MESAU	P08139 mesocricetu
4	62	28.7	2799 1 EDD_HUMAN	O05071 homo sapien
5	60.5	28.0	512 1 IE63_HSV2H	P08276 herpes simp
6	60	27.8	395 1 SGT1_YEAST	O08446 saccharomyc
7	59.5	27.5	767 1 YNM4_YEAST	P33866 saccharomyc
8	59	27.3	706 1 NUCL_MOUSE	P09405 mus musculu
9	58.5	27.1	566 1 CC45_HUMAN	O75419 homo sapien
10	58.5	27.1	582 1 MNT_HUMAN	O99583 homo sapien
11	58.5	27.1	582 1 PESC_BRARE	P79741 brachydanio
12	58	26.9	511 1 IE63_HSV1E	P36295 herpes simp
13	58	26.9	512 1 IE63_HSV1I	P20945 emericella
14	58	26.9	766 1 ABAA_EMBNI	O00363 epiliagus vi
15	58	26.9	1738 1 YCF1_EPRVI	O09536 canine oral
16	57	26.4	597 1 VCI1_COPY	O03066 xenopus lae
17	56.5	26.2	580 1 RGPI_XENILA	O13066 xenopus lae
18	56.5	26.2	692 1 ANR6_HUMAN	O3244 homo sapien
19	56	25.9	517 1 VMT2_BOVIN	O07963 bos taurus
20	56	25.9	522 1 PLIN_HUMAN	O60240 homo sapien
21	56	25.9	1711 1 CHD1_MOUSE	P40201 mus musculu
22	55.5	25.7	411 1 MP62_LYPI	P91753 lytechinus
23	55.5	25.7	618 1 ORC2_DROME	O24168 drosophilla
24	55.5	25.7	1465 1 DPOA_MOUSE	P33609 mus musculu
25	55	25.5	411 1 FKBP3_YEAST	P38911 saccharomyc
26	55	25.5	470 1 SSF1_MOUSE	O91y48 mus musculu
27	55	25.5	503 1 DMP1_MOUSE	O5128 mus musculu
28	55	25.5	732 1 NPH1_HUMAN	O51259 homo sapien
29	55	25.5	1040 1 BO12_YEAST	P39569 saccharomyc
30	55	25.5	1241 1 TRK1_SACBA	P38569 saccharomyc
31	54.5	25.2	440 1 SCAL_DROME	P30052 drosophilla
32	54.5	25.2	899 1 YMJ3_YEAST	O04500 saccharomyc
33	54	25.0	205 1 EF1B_YEAST	P32471 saccharomyc

34	54	25.0	332 1 SP2B_BACSU	P37575 bacillus su
35	54	25.0	489 1 DMP1_RAT	P81813 ratius norv
36	54	25.0	529 1 VLZ7_CAEEL	P34420 caenorhabdi
37	54	25.0	677 1 UBFL_XENILA	P25979 xenopus lae
38	54	25.0	699 1 SRCH_HUMAN	P23327 homo sapien
39	54	25.0	1085 1 IFH1_YEAST	P39520 saccharomyc
40	54	25.0	1912 1 VIT1_CHICK	P87498 gallus gall
41	54	25.0	2414 1 P300_HUMAN	O09472 homo sapien
42	53.5	24.8	891 1 MAZ3_SCHCO	P37937 schizophyil
43	53.5	24.8	1025 1 MK21_YEAST	O12176 saccharomyc
44	53.5	24.8	1108 1 CN3B_RAT	O63085 ratius norv
45	53.5	24.8	1324 1 CUF3_SCHPO	P41004 schizosacch

ALIGNMENTS

RESULT 1
ID NUCL_HUMAN STANDARD; PRT; 706 AA.
AC P19338;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89290043; PubMed=2737305;
RA Striastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT "Cloning and sequencing of the human nucleolin cDNA."
RL FEBS Lett. 250:99-105(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368666; PubMed=2394707;
RA Striastava M., McBride O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT "Genomic organization and chromosomal localization of the human
nucleolin gene."
RL J. Biol. Chem. 265:14922-14931(1990).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL: M60858; AAA59954.1; -
PIR: A35804; A35804.
DR PIR: S04631; S04631.
DR HSSP: P09651; 1HA1.
DR Aarhuus/Genet-2DPAGE: 1210; NEPHGE.
DR Genew: HGNC:7667; NCL.
DR MIM: 164035;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 4.
DR SMART: SM00360; RRM; 4.
DR PROSITE: PS50102; RRM; 4.
DR PROSITE: PS00030; RRM_RNP_1; 3.
KW Nuclear protein; phosphorylation; Methylation; DNA-binding; Repeat;
RNA-binding.

```

FT INIT_MET 0 0
CC DOMAIN 142 170 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 184 208 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 233 270 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 306 382 RNA-BINDING (RRM) 1.
FT DOMAIN 392 465 RNA-BINDING (RRM) 2.
FT DOMAIN 485 559 RNA-BINDING (RRM) 3.
FT DOMAIN 571 643 RNA-BINDING (RRM) 4.
FT DOMAIN 645 694 ARG/GLY/PHE-RICH.
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT REPEAT 57 64 X-X.
FT REPEAT 74 81 1.
FT REPEAT 82 89 2.
FT REPEAT 90 97 3.
FT REPEAT 98 103 4.
FT REPEAT 104 111 5 (INCOMPLETE).
FT REPEAT 119 126 6.
FT REPEAT 127 134 7.
FT MOD_RES 144 144 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 152 152 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 183 183 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 706 AA; 76213 MW; 85A2FC2CA22EA03DB CRC64;

Query Match 29.6%; Score 64; DB 1; Length 706;
Best Local Similarity 35.3%; Pred. No. 2.7;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSEDEDFEETAGVSPPRGKE 34
Db 186 EDEDEDEDEDEDEDEDEDESEEMETTPAKGK 219

RESULT 2
ID NUCL_RAT STANDARD; PRT; 712 AA.
AC P13383;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL OR NUC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269607; PubMed=2347493;
RA Bourdon H.-M., Amalric F.;
RT "Nucleolin gene organization in rodents: highly conserved sequences
RL within three of the 13 Introns."
RL Gene 88:187-196(1990).
RN [2]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=89121496; PubMed=2906027;
RA Bourdon H.-M., Prudhomme M., Amalric F.;
RT "Sequence and structure of the nucleolin promoter in rodents:
RL characterization of a strikingly conserved CpG island."
RL Gene 68:73-84(1988).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDICES CHROMATIN
CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55022; AAAA1732.1; -
DR EMBL; M55015; AAAA1732.1; JOINED.
DR EMBL; M55017; AAAA1732.1; JOINED.
DR EMBL; M55020; AAAA1732.1; JOINED.
DR EMBL; M25020; AAAA1732.1; JOINED.
DR PIR; JH0148; JH0148.
DR HSSP; P09651; 1HA1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; PS0102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; 3.
DR KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding.
FT INIT_MET 0 0
FT DOMAIN 142 167 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 187 215 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 241 274 ASP/GLU-RICH (ACIDIC)..
FT DOMAIN 310 386 RNA-BINDING (RRM) 1.
FT DOMAIN 396 469 RNA-BINDING (RRM) 2.
FT DOMAIN 488 562 RNA-BINDING (RRM) 3.
FT DOMAIN 574 645 RNA-BINDING (RRM) 4.
FT DOMAIN 57 702 ARG/GLY/PHE-RICH.
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT REPEAT 57 64 X-X.
FT REPEAT 74 81 1.
FT REPEAT 82 89 2.
FT REPEAT 90 97 3.
FT REPEAT 98 103 4.
FT REPEAT 104 111 5 (INCOMPLETE).
FT REPEAT 119 126 6.
FT REPEAT 127 134 7.
FT REPEAT 127 134 8.
SQ SEQUENCE 712 AA; 77016 MW; 68774A214E50F90 CRC64;

Query Match 29.6%; Score 64; DB 1; Length 712;
Best Local Similarity 29.4%; Pred. No. 2.7;
Matches 10; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSEDEDFEETAGVSPPRGKE 34
Db 193 DDDDEDEDEDEDEDEDEDESEEMETTPAKGK 226

RESULT 3
ID NUCL_MESAU STANDARD; PRT; 713 AA.
AC P08199;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE OF 35-713 FROM N.A.; AND SEQUENCE OF 1-34.
RX MEDLINE=87175501; PubMed=3470736;
RA Lapeyre B., Bourdon H., Amalric F.;
RT "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
RL an unusual protein structure revealed by the nucleotide sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
RN [2]
RP FUNCTION.
RX MEDLINE=88312631; PubMed=3409881;
RX Ertad M.S., Belenguer P., Calzergues-Ferrer M., Pantaloni A.,
RA Amalric F.;

```



```

CC -----
DR EMBL: AF006010; AAD01259.2; -
DR EMBL: U95000; AAF88143.1; -
DR EMBL: AB020703; BAA74919.1; -
DR PDB: 112T; 18-APR-02.
DR InterPro: IPR000569; HECT_domain.
DR InterPro: IPR002004; PABP/HECT.
DR InterPro: IPR000449; UBA_domain.
DR InterPro: IPR003126; znf_nrecogln.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF00632; HECT; 1.
DR Pfam: PF00658; PABP; 1.
DR Pfam: PF02207; zf-UBR1; 1.
DR SMART: SM00119; HECTC; 1.
DR SMART: SM00517; POLYA; 1.
DR SMART: SM00396; znf-UBR1; 1.
DR PROSITE: PS50237; HECT; 1.
KW Ubl conjugation pathway; Ligase; Nuclear protein; 3D-structure.
FT DOMAIN 2393 2449 PABP-LIKE.
FT 2462 2799 HECT.
FT 1986 1997 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 2036 2059 PRO-RICH.
FT DOMAIN 2329 2348 ARG/GLU-RICH (MIXED CHARGE).
FT 2366 2366 ARG/ASP-RICH (MIXED CHARGE).
FT DOMAIN 2489 2500 ASP/GLU-RICH (ACIDIC).
FT 2737 2757 PRO-RICH.
FT DOMAIN 940 945 POLY-GLU.
FT 985 985 POLY-SER.
FT DOMAIN 1528 1537 POLY-SER.
FT 1671 1681 POLY-SER.
FT DOMAIN 1762 1768 POLY-ALA.
FT BINING 2768 2768 UBIQUITIN (BY SIMILARITY).
FT 2768 2768 C->A: LOSS OF UBIQUITIN BINDING.
FT 134 134 S-> P (IN REF. 2).
FT 229 229 E-> K (IN REF. 2).
FT 258 258 S-> Y (IN REF. 2).
FT 374 375 IG-> M (IN REF. 2).
FT 772 772 D-> H (IN REF. 2).
FT 780 780 Q-> R (IN REF. 2).
FT 884 884 D-> G (IN REF. 2).
FT 1811 1811 S-> P (IN REF. 2).
FT 2144 2144 L-> H (IN REF. 2).
FT 2282 2282 K-> R (IN REF. 2).
FT 2474 2474 MISSING (IN REF. 2).
FT 2489 2489 D-> N (IN REF. 2).
SQ SEQUENCE 2799 AA; 309348 MW; 871300DB404FF561 CRC64;

Query Match 28.7%; Score 62; DB 1; Length 2799;
Best Local Similarity 29.8%; Pred. No. 21;
Matches 17; Conservative 6; Mismatches 18; Indels 16; Gaps 2;

OY 1 EDDERSTDS--SQCCSEDEDFEETAYSP-----RGEKRWRR 41
DB 1663 EDDSSDSSDSSSSQDIEQETFMLEPLERTTSSHANGAQAAPRSMQMAVR 1719

RESULT 5
ID IE63_HSV2H STANDARD; PRT; 512 AA.
AC IE63_HSV2H
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcriptional regulator IE63 (VMM63) (ICP27).
GN UL54.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OC NCBL_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; Pubmed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;

```

```

RT *Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2.*
RL J. Gen. Virol. 72:3057-3075(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RA Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY
CC GENES AND ACTS IN COMBINATION WITH ICP0 AND ICP4 AS AN ACTIVATOR
CC OF LATE GENES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HCMV UL69, AND HSV-1 57.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D10471; BAA01269.1; -
CC EMBL: 286099; CAB06702.1; -
CC PIR: J01498; MMBEXA.
KW Early protein; Transcription regulation; DNA-binding.
SQ SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;

Query Match 28.0%; Score 60.5; DB 1; Length 512;
Best Local Similarity 34.1%; Pred. No. 4.9;
Matches 14; Conservative 7; Mismatches 13; Indels 7; Gaps 1;

OY 3 DERSTDSQCCSEDEDFE-----ETAYSPPRGKER 36
DB 34 DPESDSGECSSDDEMEDPCGGCGAFAIDAIPKPPAR 74

RESULT 6
ID SGT1_YEAST STANDARD; PRT; 395 AA.
AC 008446;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SGT1 protein.
GN SGT1 OR YOR057M OR YOR29-08.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBL_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitagawa K., Connolly C., Hieter P.;
RA Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RX MEDLINE=97279235; Pubmed=9133743;
RA Valens M., Bohn C., Daigian-Fornier B., Dang V., Bolotin-Fukuhara M.;
RT "The sequence of a 54.7 Kb fragment of yeast chromosome XV reveals
RT the presence of two tRNAs and 24 new open reading frames."
RL Yeast 13:379-390(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U88830; AAB48841.1; -
CC EMBL: 270678; CAA94542.1; -
CC EMBL: 274965; CAA99250.1; -

```



```

FT DOMAIN 645 696 ARG/GLY/PHE-RICH.
FT MOD_RES 144 144 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 156 156 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 188 188 PHOSPHORYLATION (BY SIMILARITY).
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-X-X.
FT REPEAT 57 64 1.
FT REPEAT 74 81 2.
FT REPEAT 82 89 3.
FT REPEAT 90 97 4.
FT REPEAT 98 103 5 (INCOMPLETE).
FT REPEAT 104 111 6.
FT REPEAT 119 126 7.
FT REPEAT 127 134 8.
SQ SEQUENCE 706 AA: 76592 MW: 89505EE39C89F832 CRC64;

Query Match 27.3%; Score 59; DB 1; Length 706;
Best Local Similarity 32.4%; Pred. No. 11;
Matches 11; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

OY 1 EDDERSTDSQCCSEDEDFEETAGVSPRGKKE 34
Db 192 EDEDEDEDEDEDEDEDEDEEVEEMETITAKGKK 225

RESULT 9
CC45_HUMAN STANDARD; PRT; 566 AA.
ID CC45_HUMAN 060856;
AC 075419; 060856;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CDC45-related protein (PORC-PI-1).
GN CDC45L OR CDC45L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98325026; PubMed=9660782;
RA Saha P., Thome K.C., Yamaguchi R., Hou Z.-H., Werenowicz S., Dutta A.;
RT "The human homolog of Saccharomyces cerevisiae CDC45."
RL J. Biol. Chem. 273:18205-18209(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-Brain;
RA MEDLINE=99160479; PubMed=10051334;
RA Shaikh T.H., Gottlieb S., Sellinger B., Chen F., Roe B.A., Oakley R.J.,
RT Emanuel B.S., Budarf M.L.;
RT "Characterization of CDC45L: a gene in the 22q11.2 deletion region
expressed during murine and human development."
RL Mamm. Genome 10:322-326(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98391768; PubMed=9724329;
RA McKie J.M., Wadey R.B., Sutherland H.F., Taylor C.L., Scambler P.J.;
RT "Direct selection of conserved CDNs from the Digeorge critical
region: Isolation of a novel CDC45-like gene."
RL Genome Res. 8:834-841(1998).
CC -1- FUNCTION: REQUIRED FOR INITIATION OF CHROMOSOMAL DNA REPLICATION.
CC -1- SUBUNIT: ASSOCIATED WITH ORC2L.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED, HIGHEST LEVELS ARE FOUND IN
ADULT TESTIS AND TYMUS AND IN FETAL LIVER.
CC -1- DEVELOPMENTAL STAGE: TRANSCRIPT PEAKS AT G1-S TRANSITION, BUT
TOTAL PROTEIN REMAINS CONSTANT THROUGHOUT THE CELL CYCLE.
CC EXPRESSED IN MULTIPLE TISSUES DURING EMBRYOGENESIS, INCLUDING
NEURAL CREST-DERIVED STRUCTURES.
CC -1- SIMILARITY: BELONGS TO THE CDC45 FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

```

CC, the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: AF053074; AAC27289.1; -
DR EMBL: AF081535; AAD08998.1; -
DR EMBL: AJ223728; CAA11530.1; -
DR Genbank: HGNC:1739; CDC45L.
DR MIM: 603465; -
DR InterPro: IPR003874; CDC45_Like.
DR Pfam: PF02724; CDC45; 1.
DR DNA replication: Cell cycle; Nuclear protein.
FT CONFLICT 115 115 I -> V (IN REF. 3).
SQ SEQUENCE 566 AA: 65568 MW: A1BE2C8C8E5F867 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 566;
Best Local Similarity 29.8%; Pred. No. 9.6;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

OY 1 EDDERS-----TDSQCCSEDEDFEETAGVSPRGKKE 41
Db 139 EDEHSGNDSGSEPFKRTLEELVEYQTM-----RRGRREAR 180

RESULT 10
MNT_HUMAN STANDARD; PRT; 582 AA.
ID MNT_HUMAN 099583;
AC 099583;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Max binding protein MNT (ROX protein) (MYC antagonist MNT).
GN MNT OR ROX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-Fetal brain;
RA MEDLINE=97327566; PubMed=9184233;
RA Meroni G., Raymond A., Alcalay M., Borsani G., Tanigami A.,
RA Tonlorenzi R., Lo Nigro C., Messali S., Zollo M., Ledbetter D.H.,
RA Brent R., Balalbio A., Carozzo R.;
RT "Rox, a novel bHLH protein expressed in quiescent cells that
heterodimerizes with Max, binds a non-canonical E box and acts as a
transcriptional repressor."
RL EMBO J. 16:2892-2906(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98260677; PubMed=9598315;
RA Nigro C.L., Venesio T., Raymond A., Meroni G., Alberici P.,
RA Calanca S., Enrico F., Stack M., Ledbetter D.H., Lascia D.S.,
RA Balalbio A., Carozzo R.;
RT "The human ROX Gene: genomic structure and mutation analysis in human
breast tumors."
RL Genomics 49:275-282(1998).
CC -1- FUNCTION: BINDS DNA AS A HETERODIMER WITH MAX AND REPRESSSES
TRANSCRIPTION. BINDS TO THE CANONICAL E BOX SEQUENCE
5'-CACGGG-3' AND, WITH HIGHER AFFINITY, TO 5'-CACGGG-3'.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
BHLH PROTEIN. BINDS DNA AS AN HOMODIMER OR A HETERODIMER WITH MAX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X96401: CAA65265.1; -

DR EMBL: Y13440: CAA73851.1; JOINED.

DR EMBL: Y13441: CAA73851.1; JOINED.

DR EMBL: Y13442: CAA73851.1; JOINED.

DR EMBL: Y13443: CAA73851.1; JOINED.

DR EMBL: Y13444: CAA73851.1; JOINED.

DR HSSP: P25912: HLIO.

DR TRANSFAC: T03268; -

DR Genew: HGNC: 7188; MNT.

DR MIM: 603039; -

DR InterPro: IPR001092: HLH_basic.

DR Pfam: PF00010: HLH_1.

DR SMART: SM00353: HLH_1.

DR PROSITE: PS00038: HLH_1; FALSE_NEG.

DR PROSITE: PS00888: HLH_2; 1.

KW Transcription regulation; Repressor; Nuclear protein; DNA-binding.

FT DNA_BIND 222 233 BASIC DOMAIN.

FT DOMAIN 234 270 HELIX-LOOP-HELIX MOTIF (POTENTIAL).

FT DOMAIN 271 299 LEUCINE-ZIPPER.

SO SEQUENCE 582 AA; 62299 MW; 06AC320D79BF18A0 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 582;
Best Local Similarity 34.2%; Pred. No. 9.9;
Matches 13; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

OY 1 EDDERSTDSQCCSSQ-DEDFEETAOVSPRCKEKRO 37
DB 323 EDDQASTSTASGEEDNIDEMEDRAGLPKLSHRPO 360

RESULT 11

PE3C_BRARE STANDARD: PRT: 582 AA.

AC P79741; Q91906; -

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Pescadillo.

GN PES.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OC NCBI_TaxID=7955;.

OX NCBI_TaxID=7955;.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-9713157; PubMed-8985183;

RA Allende M.L., Amsterdam A., Becker T., Kawakami K., Galano N., Hopkins N.,

RT "Insertional mutagenesis in zebrafish identifies two novel genes, pescadillo and dead eye, essential for embryonic development.";

RL Genes Dev. 10:3141-3155(1996).

RN [2]

RP SEQUENCE OF 1-180 FROM N.A.

RA Kawakami K., Grosshans H., Hopkins N.,

RT "The genomic sequence of the zebrafish pescadillo gene.";

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: May have an essential role for embryonic development.

CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.

CC -1- DEVELOPMENTAL STAGE: Widely and highly expressed during the first 3 days of embryogenesis. Prominent sites of expression are the eyes and optic tectum on day 1, the fin buds, liver primordium, and gut on day 2, and the branchial arches on day 3.

CC -1- MISCELLANEOUS: PES mutant embryos exhibit smaller eyes, a reduced brain and visceral skeleton, shortened fins and a lack of expansion of liver and gut, and die on day 6 of development.

CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U77627: AAB61138.1; -

DR EMBL: AB046115: BAB01764.1; -

DR EMBL: AB046108: BAB01764.1; JOINED.

DR ZFIN: ZDB-GENE-990415-206; pes.

DR InterPro: IPR001357; BRCT.

DR Pfam: PF00533; BRCT; 1.

DR SMART: SM00292; BRCT; 1.

DR PROSITE: PS00172; BRCT; 1.

KW Nuclear protein; coiled coil.

FT DOMAIN 269 329 COILED COIL (POTENTIAL).

FT DOMAIN 323 416 BRCT.

FT DOMAIN 453 543 COILED COIL (POTENTIAL).

FT DOMAIN 537 557 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 455 470 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 472 483 ASP/GLU-RICH (ACIDIC).

SO SEQUENCE 582 AA; 67770 MW; 51ED59F03DEBF34 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 582;
Best Local Similarity 34.5%; Pred. No. 9.9;
Matches 20; Conservative 6; Mismatches 11; Indels 21; Gaps 3;

OY 1 EDDERSTDSQCCSSEDFEETAOVSPRCKEKRO 41
DB 466 EDDDEDEDEQ--SEDEEEAEANLAEMERKRSQKSLSVKVP--GKAKEENRR 519

RESULT 12

IE63_HSV1E STANDARD: PRT: 511 AA.

AC P36295; -

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Transcriptional regulator IE63 (VMM63) (ICP27).

GN US4.

OS Herpes simplex virus (type 1 / strain HPEM).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.

OC NCBI_TaxID=10303;.

OX NCBI_TaxID=10303;.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-93070559; PubMed-1332274;

RA Rosen-Wolff A., Frank S., Raab K., Moyal M., Becker Y., Darai G.,

RT "Determination of the coding capacity of the BamHI DNA fragment B of the pathogenic Herpes simplex virus type 1 strain HPEM by DNA nucleotide sequence analysis.";

RL Virus Res. 25:189-199(1992).

CC -1- FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY GENES AND ACTS IN COMBINATION WITH ICP0 AND ICP4 AS AN ACTIVATOR OF LATE GENES.

CC -1- PTM: PHOSPHORYLATED.

CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 US4, HSV-2 US54, EBV-1 5, VZV 4, EBV BMLP1, HCMV UL69, AND HVS-1 57.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M90438; -; NOT_ANNOTATED_CDS.

DR PIR: A48560; A48560.

DR PIR: A48560; A48560.

KW Early protein; Transcription regulation; DNA-binding; Phosphorylation.

SO SEQUENCE 511 AA; 55142 MW; 7BE7A8F841A98174 CRC64;


```

OS Epifagus virginiana (Beechdrops).
OG Chloroplast.
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Epifagus.
OX NCBI_TaxID=4177;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92114084; PubMed=1731088;
RA Wolfe K.H., Morden C.W., Palmer J.D.;
RT "Small single-copy region of plastid DNA in the non-photosynthetic
RT angiosperm Epifagus virginiana contains only two genes. Differences
RT among dicots, monocots and bryophytes in gene organization at a non-
RT bioenergetic locus."
RL J. Mol. Biol. 223:95-104(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9306301; PubMed=1332054;
RA Wolfe K.H., Morden C.W., Palmer J.D.;
RT "Function and evolution of a minimal plastid genome from a
RT nonphotosynthetic parasitic plant."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10648-10652(1992).
CC -!- FUNCTION: NOT YET KNOWN.
CC -!- SIMILARITY: BELONGS TO THE YCF1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X61368; CAA43644.1; -;
DR EMBL: M81884; AAA65870.1; -;
DR PIR: S16720; S16720.
DR PIR: S20614; S20614.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 1738 AA; 208318 MW; 65c63f63bdc8364b CRC64;

Query Match 26.98; Score 58; DB 1; Length 1738;
Best Local Similarity 48.48; Pred. No. 38;
Matches 15; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

QY 7 TDSOOCSEDEDIFETTAQVSPPGKRRQ 37
   | | | | | | | | | | | | | | |
Db 245 TLKKEQKVEDEDI--ETVNMPLGLKRE 273

```

Search completed: June 17, 2003, 11:56:45
 Job time : 3.4612 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:52:36 ; Search time 4.38214 Seconds
(Without alignments)
1000.321 Million cell updates/sec

Title: US-09-991-681-28
Perfect score: 216
Sequence: 1 EDDERSTDSQQCSSEDEI.....ETRAQVSPRGKRRQRRAR 41

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published_Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	29.2	714	10 US-09-978-242-3	Sequence 3, App1
2	62	28.7	268	10 US-09-864-761-33475	Sequence 4, App1
3	62	28.7	2799	9 US-10-151-736-4	Sequence 156, App1
4	60.5	28.0	512	9 US-10-121-988-156	Sequence 80, App1
5	59	27.3	957	10 US-09-801-574-80	Sequence 41, App1
6	58.5	27.1	566	9 US-09-992-598-41	Sequence 41, App1
7	58.5	27.1	566	9 US-09-989-293A-41	Sequence 41, App1
8	58.5	27.1	566	9 US-09-989-735-41	Sequence 41, App1
9	58.5	27.1	566	9 US-09-989-444-41	Sequence 41, App1
10	58.5	27.1	566	9 US-09-989-730-41	Sequence 41, App1
11	58.5	27.1	566	9 US-09-990-436-41	Sequence 41, App1
12	58.5	27.1	566	9 US-09-991-181-41	Sequence 41, App1
13	58.5	27.1	566	9 US-09-993-687-41	Sequence 41, App1
14	58.5	27.1	566	9 US-09-989-734-41	Sequence 41, App1
15	58.5	27.1	566	9 US-09-997-653-41	Sequence 41, App1
16	58.5	27.1	566	9 US-10-174-590-56	Sequence 56, App1
17	58.5	27.1	566	9 US-10-176-758-56	Sequence 56, App1
18	58.5	27.1	566	9 US-10-175-737-56	Sequence 56, App1
19	58.5	27.1	566	9 US-09-993-667-41	Sequence 41, App1

20	58.5	27.1	566	9 US-10-173-706-56	Sequence 56, App1
21	58.5	27.1	566	9 US-10-175-738-56	Sequence 56, App1
22	58.5	27.1	566	9 US-10-175-752-56	Sequence 56, App1
23	58.5	27.1	566	9 US-10-176-482-56	Sequence 56, App1
24	58.5	27.1	566	9 US-10-176-757-56	Sequence 56, App1
25	58.5	27.1	566	9 US-10-180-552-56	Sequence 56, App1
26	58.5	27.1	566	9 US-10-180-552-56	Sequence 56, App1
27	58.5	27.1	566	9 US-10-180-557-56	Sequence 56, App1
28	58.5	27.1	566	9 US-09-990-438-41	Sequence 41, App1
29	58.5	27.1	566	9 US-09-990-562-41	Sequence 41, App1
30	58.5	27.1	566	9 US-09-997-428-41	Sequence 41, App1
31	58.5	27.1	566	9 US-09-997-666-41	Sequence 41, App1
32	58.5	27.1	566	9 US-10-173-700-56	Sequence 56, App1
33	58.5	27.1	566	9 US-10-174-572-56	Sequence 56, App1
34	58.5	27.1	566	9 US-10-174-579-56	Sequence 56, App1
35	58.5	27.1	566	9 US-10-174-582-56	Sequence 56, App1
36	58.5	27.1	566	9 US-10-174-588-56	Sequence 56, App1
37	58.5	27.1	566	9 US-10-175-739-56	Sequence 56, App1
38	58.5	27.1	566	9 US-10-175-740-56	Sequence 56, App1
39	58.5	27.1	566	9 US-10-175-743-56	Sequence 56, App1
40	58.5	27.1	566	9 US-10-176-488-56	Sequence 56, App1
41	58.5	27.1	566	9 US-10-176-492-56	Sequence 56, App1
42	58.5	27.1	566	9 US-10-176-747-56	Sequence 56, App1
43	58.5	27.1	566	9 US-10-176-750-56	Sequence 56, App1
44	58.5	27.1	566	9 US-10-176-985-56	Sequence 56, App1
45	58.5	27.1	566	9 US-10-176-987-56	Sequence 56, App1

ALIGNMENTS

RESULT 1
US-09-978-242-3
Sequence 3, Application US/09978242
Patent No. US2002098566A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Yue, Henry
Corley, Neil C.
Shah, Purvi
TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/978,242
FILING DATE: 15-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/241,333
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/990,114
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PE-0451 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

```

      :          LENGTH: 714 amino acids
      :          TYPE: amino acid
      :          STRANDEDNESS: single
      :          TOPOLOGY: linear
      :          IMMEDIATE SOURCE:
      :          LIBRARY: GenBank
      :          CLONE: 128842
      :          SEQUENCE DESCRIPTION: SEQ ID NO: 3:
      :          US-09-978-242-3

Query Match           29.2%;   Score 63;   DB 10;   Length 714
Best Local Similarity 32.4%;   Pred. No. 4.9;
Matches 12; Conservative 8; Mismatches 17; Indels

Oy       1 EDDERSTSSQCCSSEDEDIFFETRNQVSPPRGKERKQ 37
         ||:::|:::||||::|||::|
Db       248 EDEDEEEDEEEDDEEEDEEEEPVKAPAGRRKE 284

```

```

1  RESULT 2
2  US-09-864-761-33475
3  Sequence 33475, Application US/09864761
4  Patent No. US20020048763A1
5  GENERAL INFORMATION:
6  APPLICANT: Penn, Sharon G.
7  APPLICANT: Rank, David R.
8  APPLICANT: Hanzel, David K.
9  APPLICANT: Chen, Wensheng
10 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
11 FILE REFERENCE: Aecomica-X-1
12 FILE REFERENCE: GENE EXPRESSION ANALYSIS BY MICROARRAY
13 CURRENT FILING DATE: 2001-05-23
14 PRIOR APPLICATION NUMBER: US/09/864,761
15 PRIOR APPLICATION NUMBER: US 60/180,312
16 PRIOR FILING DATE: 2000-02-04
17 PRIOR APPLICATION NUMBER: US 60/207,456
18 PRIOR FILING DATE: 2000-05-26
19 PRIOR APPLICATION NUMBER: US 09/632,366
20 PRIOR FILING DATE: 2000-08-03
21 PRIOR APPLICATION NUMBER: GB 24263.6
22 PRIOR FILING DATE: 2000-10-04
23 PRIOR APPLICATION NUMBER: US 60/236,359
24 PRIOR FILING DATE: 2000-09-27
25 PRIOR APPLICATION NUMBER: PCT/US01/00666
26 PRIOR FILING DATE: 2001-01-30
27 PRIOR APPLICATION NUMBER: PCT/US01/00667
28 PRIOR FILING DATE: 2001-01-30
29 PRIOR APPLICATION NUMBER: PCT/US01/00664
30 PRIOR FILING DATE: 2001-01-30
31 PRIOR APPLICATION NUMBER: PCT/US01/00669
32 PRIOR FILING DATE: 2001-01-30
33 PRIOR APPLICATION NUMBER: PCT/US01/00665
34 PRIOR FILING DATE: 2001-01-30
35 PRIOR APPLICATION NUMBER: PCT/US01/00668
36 PRIOR FILING DATE: 2001-01-30
37 PRIOR APPLICATION NUMBER: PCT/US01/00663
38 PRIOR FILING DATE: 2001-01-30
39 PRIOR APPLICATION NUMBER: PCT/US01/00662
40 PRIOR FILING DATE: 2001-01-30
41 PRIOR APPLICATION NUMBER: PCT/US01/00661
42 PRIOR FILING DATE: 2001-01-30
43 PRIOR APPLICATION NUMBER: PCT/US01/00670
44 PRIOR FILING DATE: 2001-01-30
45 PRIOR APPLICATION NUMBER: US 60/234,687
46 PRIOR FILING DATE: 2000-09-21
47 PRIOR APPLICATION NUMBER: US 09/608,408
48 PRIOR FILING DATE: 2000-06-30
49 PRIOR APPLICATION NUMBER: US 09/774,203
50 PRIOR FILING DATE: 2001-01-29
51 NUMBER OF SEQ ID NOS: 49117
52 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
53 SEQ ID NO 33475
54 LENGTH: 268

```

```

1 TYPE: PRT
2 ORGANISM: Homo sapiens
3 FEATURE:
4 OTHER INFORMATION: MAP TO AF118808.1
5 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
6 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.87
7 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
8 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
9 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
10 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
11 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
12 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.33
13 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.74
14 OTHER INFORMATION: EXPRESSED IN B4474, SIGNAL = 4.1
15 OTHER INFORMATION: EST. HUMAN HIT: BC95973.1, EVALUE 5.00e-83
16 OTHER INFORMATION: SWISSPROT HIT: P49910, EVALUE 7.00e-38
17
18 US-09-864-761-33475

```

```

QY      10 S0QCSSED--EDIFEETAQVSPPRGKEXRQMR 39
      11: 111 :11:11 :11 :11
Db      17 S0RAKSEDHVXNIFKETEEMSKTEGKLENCWR 48

RESULT 3
US-10-151-736-4
; Sequence 4, Application US/10151736
; Publication No. US20020192160A1
; GENERAL INFORMATION:
; APPLICANT: Callaghan, Michelle J.
; APPLICANT: Sutherland, Lindfield
; APPLICANT: Watts, Colin K.
; TITLE OF INVENTION: No. US20020192160A1el Human Tumour Suppressor Gene
; FILE REFERENCE: RICE-0100CN
; CURRENT APPLICATION NUMBER: US/10/151,736
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/403,402
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: PCT/AU98/00280
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ. ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2799
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-736-4

Query Match      28.7%; Score 62; DB 9; Length 2799;
Best Local Similarity 29.8%; Pred. No. 31;
Matches 17; Conservative 6; Mismatches 18; Indels 16; Gaps 2;

QY      1 EDDERSDSDS--S0QCSSEDEIFETETAQVSP-----RGKEXRQMRAR 41
      111 :11:11 :11:11 :11 :11
Db      1663 EDDSDSDSDSDSSSDSDDI EDETFMLDEPLERTNTSSHANGAQAQPSRWQWAVR 1719

RESULT 4
US-10-121-988-156
; Sequence 156, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

```

```

: TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
: FILE REFERENCE: 210121.538C1
: CURRENT APPLICATION NUMBER: US/10/121.988
: CURRENT FILING DATE: 2002-04-11
: NUMBER OF SEQ ID NOS: 183
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 156
: LENGTH: 512
: TYPE: PRT
: ORGANISM: HSV2
: US-10-121-988-156

Query Match      28.0%; Score 60.5; DB 9; Length 512;
Best Local Similarity 34.1%; Pred. No. 7.1;
Matches 14; Conservative 7; Mismatches 13; Indels 7; Gaps 1;

OY      3 DEDSTSSQCCSEDEDFE-----ETNAVSPPRKEKR 36
Db      34 DDEPSDSSGECSSDEDEMDPCGDGAGAIIDAIKGPAPR 74

RESULT 5
US-09-801-574-80
: Sequence 80, Application US/09801574
: Patent No. US20020081592A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Peijng Jeremy
: APPLICANT: Page, David C.
: TITLE OF INVENTION: Reproduction-Specific Genes
: FILE REFERENCE: 0399.2007-002
: CURRENT APPLICATION NUMBER: US/09/801.574
: CURRENT FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: 60/187,518
: PRIOR FILING DATE: 2000-03-07
: PRIOR APPLICATION NUMBER: 60/261,557
: PRIOR FILING DATE: 2001-01-12
: NUMBER OF SEQ ID NOS: 90
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 80
: LENGTH: 957
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-801-574-80

Query Match      27.3%; Score 59; DB 10; Length 957;
Best Local Similarity 31.6%; Pred. No. 23;
Matches 12; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

OY      1 EDDERSTDSSQCCSEDEDFEETAYQVSPPRKEKROW 38
Db      836 ETDCKKEDSSMLSKETEDGEDGERAHSTLDELELRN 873

RESULT 6
US-09-992-598-41
: Sequence 41, Application US/09992598
: Patent No. US20020160384A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Bolstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gottlsen, Mary E.
: APPLICANT: Godowski, Audrey
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Kijavich, Ivar J.
: APPLICANT: Napier, Mary A.

APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2730PIC20
: CURRENT APPLICATION NUMBER: US/09/992.598
: CURRENT FILING DATE: 2001-11-14
: PRIOR APPLICATION NUMBER: 60/049787
: PRIOR FILING DATE: 1997-06-16
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/075945
: PRIOR FILING DATE: 1998-02-25
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087609
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087827
: PRIOR FILING DATE: 1998-06-03
: PRIOR APPLICATION NUMBER: 60/088021
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088025
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088028
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088029
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088030
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088033
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088326
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088167
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088202
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088212
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088217
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088655
: PRIOR FILING DATE: 1998-06-09
: PRIOR APPLICATION NUMBER: 60/088734
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088738
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088742
```

1	PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/0888810
2	PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/0888824
3	PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/0888826
4	PRIOR FILING DATE: 1998-06-10	PRIOR APPLICATION NUMBER: 60/0888858
5	PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/0888861
6	PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/088876
7	PRIOR FILING DATE: 1998-06-11	PRIOR APPLICATION NUMBER: 60/089105
8	PRIOR FILING DATE: 1998-06-12	PRIOR APPLICATION NUMBER: 60/089440
9	PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089512
10	PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089514
11	PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089532
12	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089558
13	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089580
14	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089599
15	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089600
16	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089653
17	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089801
18	PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089907
19	PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089908
20	PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089947
21	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089948
22	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089952
23	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/090254
24	PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090435
25	PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090444
26	PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090445
27	PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090429
28	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090431
29	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090435
30	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090444
31	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090445
32	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090472
33	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090535
34	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090540
35	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090542
36	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090544

PRIOR APPLICATION DATE: 1998-06-24	NUMBER: 60/090557
PRIOR APPLICATION NUMBER: 60/090676	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090678	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090690	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090694	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090695	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090696	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090862	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/090863	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/091366	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091478	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091544	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091519	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091626	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091633	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091978	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/091982	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/092182	
PRIOR FILING DATE: 1998-07-09	

Query Match	27.1%	Score 58.5;	DB 9;	Length 566;
Best Local Similarity	29.8%	Pred. No. 15;		
Matches 14; Conservative 13; Mismatches 9;				Gaps 2

```

QY      1 EDDERS-----TSSQCCSSEDEDIFETTAQVSPPRKEKROWRAR 41
      ||:| |:::|:| |:| |
DB      139 EDEHSGNDSGSESEKTRLEELIYEQTM-----RRQREWEAR 180

```

RESULT 7
US-09-989-293A-41

; Sequence 41, Application US/09989293A
; Patent No. US20020177164A1

; GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

• APPLICANT: Desnoyers, Luc

APPLICANT: Falcon, Dan L.

APPLICANT: Feriala, Napoleone

APPLICANT: Fong, Silelmai

APPLICANT: GEIDEL, Hanspeter

APPLICANT: Goddard Andrew

APPLICANT: Godowski Paul

APPLICANT: Grimaldi, J Christ

APPLICANT: Gurney Austin

APPLICANT: K1 Javlin, Ivar J

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas

APPLICANT: ROY, Margaret A

APPLICANT: Stewart, Timothy

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Collin

APPLICANT: Williams, P. M.

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690

;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 27.18; Score 58.5; DB 9; Length 565;
Best Local Similarity 29.88; Pred. No. 15;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDERS-----TDSQOCSEDEDIFFETAGVSPRGKRRQWRR 41
Db 139 EDEHSGNDSDGSEPSKRRLEEEIYEQTM-----RRRRRRREAR 180

RESULT 8
US-09-989-735-41
; Sequence 41, Application US/09989735
; Publication No. US20020193299A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC61
; CURRENT APPLICATION NUMBER: US/09/989,735
; CURRENT FILING DATE: 2001-11-19

;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105

;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 15;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

Qy 1 EDDERS-----TDSOCCSSEDEDFEFTQVSPRKEKROWAR 41
Db 139 EDEHSGNDSGSEPSERKTRLEERIVBQTM----RRORREWEAR 180

RESULT 9
US-09-990-444-41
; Sequence 41, Application US/09990444
; Publication No. US20020193300A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bostein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC19
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13	PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945	PRIOR FILING DATE: 1998-02-25
PRIOR FILING DATE: 1998-03-20	PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106	PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759	PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827	PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088212	PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555	PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088734	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826	PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876	PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105	PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440	PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512	PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514	PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089535323	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08953538	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08953598	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08953999	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08956000	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08956533	PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/08958001	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/0895948	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/08959322	PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/090246	PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/0903555	PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/0904299	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0904444	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0904455	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0904722	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0905355	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0905400	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0905422	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/0905577	PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090624	PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/0906594	PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/0906595	PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/0906636	PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/0908626	PRIOR FILING DATE: 1998-06-26	PRIOR APPLICATION NUMBER: 60/0908636	PRIOR FILING DATE: 1998-06-26	PRIOR APPLICATION NUMBER: 60/0913600	PRIOR FILING DATE: 1998-07-01	PRIOR APPLICATION NUMBER: 60/091478	PRIOR FILING DATE: 1998-07-01
--	-------------------------------	---------------------------------------	-------------------------------	---------------------------------------	-------------------------------	---------------------------------------	-------------------------------	---------------------------------------	-------------------------------	---------------------------------------	-------------------------------	---------------------------------------	-------------------------------	--------------------------------------	-------------------------------	---------------------------------------	-------------------------------	-------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	-------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	--------------------------------------	-------------------------------	-------------------------------------	-------------------------------

PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 15;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDERS-----TDSQCCSEDEDFEETAQVSPPRGKRRAR 41
Db 139 EDEHSGNDSGSGSEKRRLEBEIVEQTM-----RRORREWEAR 180

RESULT 10
US-09-989-730-41
Sequence 41, Application US/09989730
Publication No. US20020197674A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C69
CURRENT APPLICATION NUMBER: US/09/989,730
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599

[illegible]

Publication No. US20030003531A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavits, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C64
CURRENT FILING DATE: US/09/989,734
PRIOR APPLICATION NUMBER: 2001-11-19
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429

```
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      27.1%  Score 58.5;  DB 9;  Length 566;
Best local Similarity 29.8%;  Pred. No. 15;
Matches 14;  Conservative 13;  Mismatches 9;  Indels 11;  Gaps 2;

OY      1  EDDERS-----TDSQOCSSEDEDIFFETAYVSPRGKKEKROWRAR 41
Db      139  EDEHSGNDSGSEPEKRTREEETVEJQTM-----RRQRREWEAR 180

RESULT 15
US-09-997-653-41
; Sequence 41, Application US/09997653
; Publication No. US2003008297A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
```

```
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC38
; CURRENT APPLICATION NUMBER: US/09/997,653
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
```

;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 15;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;
QY 1 EDDERS-----TDSOCCSSEDEDIFETNOVSPPRCKEKRRRAR 41
DB 139 EDERHSGNDSDGSEPSERKTRLEETVQTM-----RRORREWEAR 180

Search completed: June 17, 2003, 12:02:28
Job time : 5.38214 secs

THIS PAGE BLANK (use 70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:56:11 ; Search time 3.78184 Seconds
(without alignments)
318.982 Million cell updates/sec

Title: US-09-991-681-28
Perfect score: 216
Sequence: 1 EDDERSTDSQQCSSEDEDI.....ETNAQVSPPRCKEKRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2-6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2-6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2-6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2-6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2-6/ptodata/1/1aa/PCtUS.COMB.pep.*
- 6: /cgn2-6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	41	US-09-065-383-28	Sequence 28, Appl
2	216	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	63	29.2	714	US-08-990-114-3	Sequence 3, Appl
4	63	29.2	714	US-09-241-333-3	Sequence 3, Appl
5	57	26.4	667	US-09-303-064-55	Sequence 55, Appl
6	57	26.4	667	US-09-086-503-55	Sequence 55, Appl
7	56	25.9	421	US-08-132-649-6	Sequence 6, Appl
8	56	25.9	421	US-08-767-579-6	Sequence 6, Appl
9	56	25.9	522	US-09-232-160-23	Sequence 23, Appl
10	55.5	25.7	418	US-08-615-170-16	Sequence 16, Appl
11	55.5	25.5	180	US-08-630-915A-196	Sequence 196, App
12	55	25.5	411	US-08-741-134-6	Sequence 26, Appl
13	54	25.0	226	US-08-431-080-26	Sequence 26, Appl
14	54	25.0	226	US-08-938-534-26	Sequence 26, Appl
15	54	25.0	226	US-09-345-294-26	Sequence 26, Appl
16	54	25.0	1085	US-08-431-080-28	Sequence 28, Appl
17	54	25.0	1085	US-08-938-534-28	Sequence 28, Appl
18	54	25.0	1085	US-09-345-294-28	Sequence 28, Appl
19	54	25.0	2414	US-08-227-536-2	Sequence 2, Appl
20	54	25.0	2414	PCT-US95-04682-2	Sequence 2, Appl
21	52.5	24.3	193	US-08-679-765-3	Sequence 3, Appl
22	52.5	24.3	193	US-09-196-525-3	Sequence 3, Appl
23	52.5	24.3	193	US-09-318-317-3	Sequence 3, Appl
24	52	24.1	710	US-09-079-812E-2	Sequence 2, Appl
25	51.5	23.8	611	US-09-370-807-2	Sequence 2, Appl
26	51.5	23.8	611	US-09-921-259-2	Sequence 2, Appl
27	51.5	23.8	2476	US-08-276-967-2	Sequence 2, Appl

28	51	23.6	262	4	US-09-134-001C-3237	Sequence 3237, Ap
29	50.5	23.4	392	3	US-08-301-162-2	Sequence 2, Appl
30	50.5	23.4	392	4	US-09-461-240-2	Sequence 2, Appl
31	50.5	23.4	392	4	US-09-968-927-2	Sequence 2, Appl
32	50.5	23.4	426	1	US-08-615-170-6	Sequence 6, Appl
33	50.5	23.4	426	1	US-08-615-170-15	Sequence 15, Appl
34	50.5	23.4	428	3	US-08-301-162-18	Sequence 18, Appl
35	50.5	23.4	428	4	US-09-461-240-18	Sequence 18, Appl
36	50.5	23.4	428	4	US-09-968-927-18	Sequence 18, Appl
37	50.5	23.4	816	1	US-08-190-802A-54	Sequence 54, Appl
38	50.5	23.4	816	4	US-08-477-346-54	Sequence 54, Appl
39	50.5	23.4	816	4	US-08-473-089-54	Sequence 54, Appl
40	50.5	23.4	816	4	US-08-487-072A-54	Sequence 54, Appl
41	50	23.1	172	3	US-09-383-080-3	Sequence 3, Appl
42	50	23.1	172	3	US-08-858-937-3	Sequence 3, Appl
43	50	23.1	182	1	US-08-466-603-5	Sequence 5, Appl
44	50	23.1	182	1	US-08-314-503A-5	Sequence 5, Appl
45	50	23.1	182	1	US-08-468-066-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-28
Sequence 28, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065.383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842.385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids


```

; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-303-064-55

Query Match          26.4%; Score 57; DB 4; Length 667;
Best Local Similarity 38.9%; Pred. No. 13;
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

Oy      1 EDDERSTDSSQCCSDEDEDIREETAAOVSPPGKREKR 36
       | : || | :|||:| : | : | : | : | : | : |
Db      265 EGEGSTTESASENSEDDTTHDALQLPELGDLEVR 300

RESULT 6
US-09-086-503-55
; Sequence 55, Application US/09086503A
; Patent No. 6329157
; GENERAL INFORMATION:
; APPLICANT: MAINE, Gregory T.
; APPLICANT: HUNT, Jeffery C.
; APPLICANT: BROJANAC, Susan
; APPLICANT: JYH-TSING SHEU, Michael
; APPLICANT: CHOVAN, Linda E.
; APPLICANT: TYNER, Joan D.
; APPLICANT: HOWARD, Lawrence V.
TITLE OF INVENTION: ANTIGEN COCKTAILS AND USES THEREOF
FILE REFERENCE: 6361.US.01
CURRENT APPLICATION NUMBER: US/09/086.503A
CURRENT FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 55
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 55
LENGTH: 667
TYPE: PRT
ORGANISM: Toxoplasma gondii
US-09-086-503-55

Query Match          26.4%; Score 57; DB 4; Length 667;
Best Local Similarity 38.9%; Pred. No. 13;
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

Oy      1 EDDERSTDSSQCCSDEDEDIREETAAOVSPPGKREKR 36
       | : || | :|||:| : | : | : | : | : | : |
Db      265 EGEGSTTESASENSEDDTTHDALQLPELGDLEVR 300

RESULT 7
US-08-132-649-6
; Sequence 6, Application US/08132649
; Patent No. 5585462
; GENERAL INFORMATION:
; APPLICANT: LONDOS, Constantine
; APPLICANT: GREENBERG, Andrew S.
; APPLICANT: KIMMEL, Alan R.
; APPLICANT: EGAN, John J.
TITLE OF INVENTION: CLONING OF PERILIPIN PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Khourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0 Version #1.25

```


PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-16

Query Match 25.7%; Score 55.5; DB 1; Length 418;
Best Local Similarity 31.8%; Pred. No. 12;
Matches 14; Conservative 6; Mismatches 15; Indels 9; Gaps 1;

Qy 1 EDDERSTDSQCSSEDEFEETFAVSPPRGKER 35
Db 14 EDIERMSDADKPIDNDAGVWSPDIEGFOALAIYPCGRK 57

RESULT 11
US-08-630-915A-196
Sequence 196, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLER, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 196:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-630-915A-196

Query Match 25.5%; Score 55; DB 4; Length 180;
Best Local Similarity 28.2%; Pred. No. 5.2;
Matches 11; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

Qy 2 DDERSTDSQCSSEDEFEETFAVSPPRGKEROMRA 40
Db 114 EEEGGQSSSEGGSEDEYAVDETADGAEVKORDTPHMSA 152

RESULT 12
US-08-741-134-6
Sequence 6, Application US/08741134
Patent No. 5861498
GENERAL INFORMATION:
APPLICANT: Litwack, Gerald
APPLICANT: Alnemri, Emad S.
APPLICANT: Fernandes-Alnemri, Teresa
TITLE OF INVENTION: IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
TITLE OF INVENTION: AND
TITLE OF INVENTION: METHODS OF USING THE SAME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498-15
STREET: One Liberty Place - 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: WordPerfect for Windows 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741,134
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,163
FILING DATE: 01-NOV-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TUV-2090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-741-134-6

Query Match 25.5%; Score 55; DB 2; Length 411;
Best Local Similarity 29.7%; Pred. No. 13;
Matches 11; Conservative 9; Mismatches 15; Indels 2; Gaps 1;

Qy 1 EDDERSTDSQCSSEDEFEETFAVSPPRGKER 37
Db 221 EDEEDNDGEEDEEBE--EQKEVAKPEPKSKKE 255

RESULT 13
US-08-431-080-26
Sequence 26, Application US/08431080
Patent No. 5698686
GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-26

Query Match 25.0%; Score 54; DB 1; Length 226;
Best Local Similarity 35.7%; Pred. No. 9.1;
Matches 15; Conservative 8; Mismatches 11; Indels 8; Gaps 2;

QY 1 EDDERSTDSQCCSEDE--DIEETAAVSPPRGKRRQWRA 40
:11 :11 :11:11 :111 :111 :1
Db 146 DDGSDSDSETSSDDENIDFVKLTAA-----RKRKRAMKA 181

RESULT 14
US-08-938-534-26
Sequence 26, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,534
FILING DATE: 26-SEP-1997

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-534-26

Query Match 25.0%; Score 54; DB 2; Length 226;
Best Local Similarity 35.7%; Pred. No. 9.1;
Matches 15; Conservative 8; Mismatches 11; Indels 8; Gaps 2;

QY 1 EDDERSTDSQCCSEDE--DIEETAAVSPPRGKRRQWRA 40
:11 :11 :11:11 :111 :111 :1
Db 146 DDGSDSDSETSSDDENIDFVKLTAA-----RKRKRAMKA 181

RESULT 15
US-09-345-294-26
Sequence 26, Application US/09345294
Patent No. 6387619
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/345,294
FILING DATE: 30-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-345-294-26

Query Match 25.08; Score 54; DB 4; Length 226;
Best Local Similarity 35.7%; Pred. No. 9.1;
Matches 15; Conservative 8; Mismatches 11; Indels 8; Gaps 2;

OY 1 EDDERSTDSQCCSSEDE--DIFETTAQVSPPRGKEKROWRA 40
Db 146 DDDGSDSDSDSETSSDDENIDFVKLTIAQ-----RKRRAKMA 181

Search completed: June 17, 2003, 12:03:38
Job time : 5.78185 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:16 ; Search time 11.8258 Seconds
(without alignments)
461.981 Million cell updates/sec

Title: US-09-991-681-28
Perfect score: 216
Sequence: 1 EDERSTDSQQCSSEDEDI.....ETTAQVSPPGKEKQWRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_101002.*
1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	41	23 AAM50810	PS118 prostate mar
2	216	100.0	518	19 AAM5472	PS118 protein enco
3	216	100.0	518	23 AAM50809	PS118 prostate mar
4	216	100.0	1807	22 ABG09728	Novel human diagno
5	216	100.0	1982	22 ABG09731	Novel human diagno
6	69	31.9	192	20 AAY48248	Human prostate can
7	65.5	30.3	512	23 AAE19559	Herpes simplex vit
8	65.5	30.3	1585	22 ABG06075	Novel human diagno
9	65.5	30.3	1585	22 ABG10815	Novel human diagno
10	65.5	30.3	1598	22 ABG06618	Novel human diagno

11	65.5	30.3	1598	22 ABG09655	Novel human diagno
12	65.5	30.3	1647	22 ABG04001	Novel human diagno
13	64	29.6	536	22 ABG92836	Human protein sequ
14	64	29.6	707	16 AAR79912	Human nucleolin.
15	64	29.6	707	20 AAM84052	Human V3 loop HIV
16	64	29.6	707	22 AAB48054	Human nucleolin.
17	62.5	28.9	642	22 ABG60513	Drosophila melanog
18	62	28.7	268	22 ABG27527	Human peptide #178
19	62	28.7	268	22 ABG32686	Peptide #192 encod
20	62	28.7	268	22 ABB18177	Protein #176 encod
21	62	28.7	268	22 AAM53508	Human brain expres
22	62	28.7	268	22 AAM65886	Human bone marrow
23	62	28.7	268	22 AAM13753	Peptide #187 encod
24	62	28.7	268	22 AAM26151	Peptide #188 encod
25	62	28.7	268	22 AAM01498	Peptide #180 encod
26	62	28.7	268	23 ABG35521	Human peptide enco
27	62	28.7	2799	19 AAM81867	Human tumour suppl
28	60.5	28.0	512	15 AAR47811	HSV-2 viral protel
29	60.5	28.0	512	19 AAM73139	HSV-2 strain SB5 C
30	60.5	28.0	584	19 AAM72152	HSV-2 strain SB5 C
31	59	27.3	635	22 ABG22101	Novel human diagno
32	59	27.3	957	22 AAU07894	Polypeptide sequen
33	59	27.3	1428	22 AAG67359	Amino acid sequen
34	58.5	27.1	214	22 AAB63963	Human prostate can
35	58.5	27.1	517	23 AAU99324	Human CD63 antigen
36	58.5	27.1	566	21 AAY66641	Membrane-bound pro
37	58.5	27.1	566	22 AAU29051	Human PRO710 (UNQ3
38	58.5	27.1	566	22 AAB65164	Human brain expres
39	58	26.9	117	22 ABB42701	Peptide #10207 enc
40	58	26.9	117	22 AAM63593	Human bone marrow
41	58	26.9	117	22 AAM76408	Peptide #10555 enc
42	58	26.9	117	23 AAE19558	Herpes simplex vit
43	58	26.9	774	23 AAE14512	GST-Herpes simplex
44	58	26.9	774	23 AAE14512	Novel human diagno
45	57.5	26.6	108	22 ABG28838	

ALIGNMENTS

RESULT 1	
AAM50810	
ID	AAM50810 standard; Protein: 41 AA.
AC	AAM50810;
XX	
DT	01-MAY-2002 (first entry)
XX	
DE	PS118 prostate marker immunogenic polypeptide.
XX	
KW	PS118: prostate; marker: prostate cancer; tumour; metastasis;
KW	benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW	prostatitis; human; diagnosis; therapy; vaccine; immunogen.
XX	
OS	Homo sapiens.
XX	
PN	US2001055758-A1.
XX	
PD	27-DEC-2001.
XX	
PF	23-APR-1998; 98US-0065383.
XX	
PR	23-APR-1997; 97US-0842385.
XX	
PA	(BIL/) BILLING-MEDEL P A.
PA	(COHE/) COHEN M.
PA	(COPL/) COPLITTIS T L.
PA	(FRIE/) FRIEDMAN P N.
PA	(GORD/) GORDON J.
PA	(GRAN/) GRANADOS E N.
PA	(HODG/) HODGES S C.
PA	(KLAS/) KLAS M R.
PA	(KRAT/) KRATOCHVIL J D.

PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvill JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 43; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 184-224 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing, and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis, and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 41 AA;
 Query Match 100.0%; Score 216; DB 23; Length 41;
 Best Local Similarity 100.0%; Pred. No. 8.9e-22;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDDERSTDSQCCSSEDEDFEETAGVSPRGKEKRWRR 41
 1 EDDERSTDSQCCSSEDEDFEETAGVSPRGKEKRWRR 41
 Db
 RESULT 2
 AAM85472
 ID AAM85472 standard; Protein; 518 AA.
 XX
 AC AAM85472;
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KM EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvill JD;
 PI Robert-Strapp L, Russell JC, Stroupe SD;

XX
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.
 XX
 PT New P118 nucleic acid and proteins - used for diagnosis and
 PT treatment of prostatic disease, especially cancer, and also for drug
 PT screening
 XX
 PS Claim 17; Page 93-94; 117pp; English.
 XX
 CC The present sequence is encoded by consensus PS118 sequence derived from
 CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
 CC were identified from a cDNA library made from prostate tumour tissue.
 CC Recombinant PS118 protein is used to detect PS118-specific antibodies,
 CC to raise antibodies for detection of PS118 antigens, to screen for
 CC specific binding agents (potential therapeutics), and to isolate specific
 CC antibodies from serum. Detection of PS118 protein or nucleic acid, which
 CC are prostate related, and altered or elevated in prostatic disease, is
 CC used for detection, diagnosis, staging, monitoring and prognosis of
 CC prostatic disease, particularly cancer, and to identify subjects at
 CC risk.
 XX
 SQ Sequence 518 AA;
 Query Match 100.0%; Score 216; DB 19; Length 518;
 Best Local Similarity 100.0%; Pred. No. 1.7e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EDDERSTDSQCCSSEDEDFEETAGVSPRGKEKRWRR 41
 184 EDDERSTDSQCCSSEDEDFEETAGVSPRGKEKRWRR 224
 Db
 RESULT 3
 AAM50809
 ID AAM50809 standard; Protein; 518 AA.
 XX
 AC AAM50809;
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker partial sequence.
 XX
 KM PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KM benign prostatic hyperplasia; prostatic interepithelial neoplasia;
 KM prostatitis; human; diagnosis; therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVILL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvill JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 DR WPI; 2002-187683/24.

DR N-PSDB; ABA91651.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 42-43; 57pp; English.
 XX
 CC The present sequence is that of a human prostate-specific PS118
 CC polypeptide, as predicted from a partial consensus cDNA sequence
 CC (see ABA91651), and lacking the N-terminal region. The PS118
 CC consensus sequence is found at least 12 times more often in
 CC prostate than in non-prostate tissue. PS118 polypeptides,
 CC including derivatives of the present sequence, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells. The methods
 CC and reagents of the invention may provide an early means of
 CC detecting diseases of the prostate and may also provide new markers
 CC which can differentiate between the clinically important and
 CC unimportant prostate cancers without the use of surgery.
 CC
 SQ Sequence 518 AA:

Query Match 100.0%; Score 216; DB 23; Length 518;
 Best Local Similarity 100.0%; Pred. No. 1.7e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDDERSTDSQCCSSEDEDFEETAOVSPRGKRRKRRAR 41
 Db 184 EDDERSTDSQCCSSEDEDFEETAOVSPRGKRRKRRAR 224

RESULT 4
 ID ABOG9728 standard; Protein: 1807 AA.

XX ABOG9728;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9719.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS73915.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 40087; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABOG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 1807 AA:

Query Match 100.0%; Score 216; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 7.1e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDDERSTDSQCCSSEDEDFEETAOVSPRGKRRKRRAR 41
 Db 1473 EDDERSTDSQCCSSEDEDFEETAOVSPRGKRRKRRAR 1513

RESULT 5
 ID ABOG9731 standard; Protein: 1982 AA.

XX ABOG9731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS73918.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20; SEQ ID No 40090; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes.
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SO Sequence 1982 AA;

Query Match 100.0%; Score 216; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 7.9e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTSSQCCSFDEDFEETAAVSPRGEKRWRR 41

DB 1587 EDDERSTSSQCCSFDEDFEETAAVSPRGEKRWRR 1627

RESULT 6

AA48248
 ID AAY48248 standard; Protein; 192 AA.

AC AAY48248;

DT 08-DEC-1999 (first entry)

XX Human prostate cancer-associated protein 34.

KW Expressed sequence tag; EST; prostate tumor; antitumor; treatment;

KM gene therapy; tissue specificity human.

XX Homo sapiens.

OS

PN DE19811193-A1.

PD 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011193.

PR 10-MAR-1998; 98DE-1011193.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinemann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

XX WPI; 1999-519628/44.

PT N-PSDB; AA233451.

XX New nucleic acid expressed at high level in prostatic tumor tissue and

XX encoded polypeptides, useful for treating cancer and screening for

XX therapeutic agents.

XX Claim 22; 128; 166pp; German.

XX This invention describes novel nucleic acid sequences (A) that are

CC expressed at high level in prostatic tumor tissue and encode gene

CC products or their fragments. The products of the invention have
 CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
 CC identifying agents for treatment of prostatic cancer and (ii) for
 CC therapy of prostate cancer, optionally where expressed by gene therapy
 CC methods. (A) is also used to isolate full-length genes (for gene therapy)
 CC and for recombinant production of (I), which can be used to raise
 CC specific antibodies. (A) are identified by assembly of ESTs (expressed
 CC sequence tags) before they are analyzed for expression pattern (tissue
 CC specificity). This approach eliminates many of the false results, as
 CC regards tissue specificity, associated with known methods that use
 CC single (usually short) ESTs. AAY48215-48303 represent protein fragments
 CC encoded by the expressed sequence tags described in the method of the
 CC invention.
 XX

SO Sequence 192 AA;

Query Match 31.9%; Score 69; DB 20; Length 192;
 Best Local Similarity 100.0%; Pred. No. 0.29;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTSSQCC 13

DB 180 EDDERSTSSQCC 192

RESULT 7

AAE19559
 ID AAE19559 standard; protein; 512 AA.

AC AAE19559;

DT 31-MAY-2002 (first entry)

XX Herpes simplex virus type 2 (HSV-2) IE63 protein homologue.

KW Herpes virus infection; detection; therapy; IE63 protein; antiviral.

XX Herpes simplex virus type 2.

OS

PN WO200204492-A2.

PD 17-JAN-2002.

PF 11-JUL-2001; 2001WO-GB03114.

PR 11-JUL-2000; 2000GB-0016890.

PA (UNIU) UNIV GLASGOW.

PI Clements JB, Maclean AR;

XX WPI; 2002-226983/28.

PT Detecting an agent useful for treating herpes virus infection comprises

PT determining any change in a polypeptide/zinc complex in the presence of

PT the test agent.

XX Claim 6; Fig 1; 43pp; English.

XX The invention relates to a method of detecting an agent for use in

XX the treatment of herpes virus infection. The method comprises forming

XX a herpes virus polypeptide/zinc complex; adding a test agent to the

XX polypeptide/zinc complex; and detecting any change in the complex.

XX The invention also relates to the use of known agents, such as 2,2'-

XX diethoxybenzamide (DEBA) and azodicarbonamide (ADA), and unknown

XX agents for the manufacture of a medicament for the treatment of

XX herpes virus infections. The method is useful for detecting agents

XX for use in the treatment of herpes virus infection. The present

XX sequence is Herpes simplex virus type 2 (HSV2) IE63 protein homologue.

XX HSV2 belongs to alpha herpes virus sub-family. IE63 protein is also

XX called ICP27.

XX Sequence 512 AA;

Query Match 30.3%; Score 65.5; DB 23; Length 512;
Best Local Similarity 36.6%; Pred. No. 2.6;
Matches 15; Conservative 7; Mismatches 12; Indels 7; Gaps 1;

QY 3 DERSTDSQCCSEDEDIFE-----ETAAVSPRRCKERR 36
1: ||| ||||| : : : : :
Db 34 DPESDSDSCSESEDEMDPCGCGAIDAIAIPKPPAR 74

RESULT 8
ABG06075
ID ABG06075 standard; Protein; 1585 AA.

AC ABG06075;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6066.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS70262.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID NO 36434; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 1585 AA;

Query Match 30.3%; Score 65.5; DB 22; Length 1585;

Best Local Similarity 35.4%; Pred. No. 9.7;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

QY 1 EDERSTDSQCCS-----SEDEDIFEETAAVSPRRCKERR 41
1: ||| ||||| : : : : :
Db 1248 EENDENSLSSSDCSSENKDEETSESDIEETKEVEKEPELQTRMEER 1295

RESULT 9
ABG10815
ID ABG10815 standard; Protein; 1585 AA.

AC ABG10815;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #10806.

KM Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS75002.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID NO 41174; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 1585 AA;

Query Match 30.3%; Score 65.5; DB 22; Length 1585;
Best Local Similarity 35.4%; Pred. No. 9.7;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

QY 1 EDDERSTDSQOCS-----SEDEIFETTAQVSPRGKCKQWRRAR 41
1248 ENDENLSSSDSCSENKDEISESDIEKTEVKEPELQTRREMER 1295

RESULT 10

ABG06618
ID ABG06618 standard; Protein; 1598 AA.

AC ABG06618;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #6609.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS70805.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 36977; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1598 AA;

QY Query Match 30.3%; Score 65.5; DB 22; Length 1598;
Best Local Similarity 35.4%; Pred. No. 9.8;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

QY 1 EDDERSTDSQOCS-----SEDEIFETTAQVSPRGKCKQWRRAR 41

Db 1248 ENDENLSSSDSCSENKDEISESDIEKTEVKEPELQTRREMER 1295

RESULT 11

ABG09655
ID ABG09655 standard; Protein; 1598 AA.

AC ABG09655;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9646.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS73842.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 40014; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 1598 AA;

QY Query Match 30.3%; Score 65.5; DB 22; Length 1598;
Best Local Similarity 35.4%; Pred. No. 9.8;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

QY 1 EDDERSTDSQOCS-----SEDEIFETTAQVSPRGKCKQWRRAR 41

Db 1248 ENDENLSSSDSCSENKDEISESDIEKTEVKEPELQTRREMER 1295

RESULT 12
ID ABO04001 standard; Protein: 1647 AA.
XX
AC ABO04001;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3992.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN MO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS68188.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 20; SEQ ID NO 34360; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABO0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1647 AA;
XX
Query Match 30.3%; Score 65.5; DB 22; Length 1647;
Best Local Similarity 35.4%; Pred. No. 10;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

RESULT 13
ID AAB92836 standard; Protein: 536 AA.
XX
AC AAB92836;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11383.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 98JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs
XX
XX
PS Claim 8; SEQ ID 11383; 2537pp + CD ROM; English.
XX
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 536 AA;
XX
Query Match 29.6%; Score 64; DB 22; Length 536;
Best Local Similarity 35.3%; Pred. No. 4.4;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

```

RESULT 14
AAR79912
ID AAR79912 standard; Protein; 707 AA.
XX
AC AAR79912:
XX
DT 15-MAY-1996 (first entry)
XX
DE Human nucleolin.
XX
KM nucleolin; monoclonal antibody; immunosuppressant.
XX
OS Homo sapiens.
XX
PN JP07242566-A.
XX
PD 19-SEP-1995.
XX
PF 01-MAR-1994; 94JP-0031668.
XX
PR 01-MAR-1994; 94JP-0031668.
XX
PA (KIRI ) KIRIN BREWERY KK.
XX
DR WPI; 1995-355193/46.
XX
PT An immuno:suppressant contg. human nucleolin antibody - is effective
PT against auto:immune disease
XX
PS Disclosure; Page 9-11; 13pp; Japanese.
XX
CC The protein is that of human nucleolin. An immunosuppressant contg. an
CC antibody having reactivity to human nucleolin as the active component is
CC claimed. The immunosuppressant is effective for various immune disease
CC and is very useful in industry.
XX
SQ Sequence 707 AA:
XX
Query Match 29.6%; Score 64; DB 16; Length 707;
Best Local Similarity 35.3%; Pred. No. 6.1;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;
QY 1 EDDERSTDSSQCCSSEDEDFEETAQVSPRGKE 34
DB 187 EDDDEDDDDDDDDDEDDSEEAETTPAKGKK 220

RESULT 15
AAM84052
ID AAM84052 standard; Protein; 707 AA.
XX
AC AAM84052:
XX
DT 15-MAR-1999 (first entry)
XX
DE Human V3 loop HIV receptor p95/nucleolin protein.
XX
KM HIV receptor; V3 loop; human immunodeficiency virus; retrovirus;
KM p95 protein; nucleolin; infection; therapy; diagnosis.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 22..44 /note- "Glu/Asp-rich region, useful as inhibitor
XX FT /note- "of HIV binding (claim 6)"
XX FT 143..171 /note- "Glu/Asp-rich region, useful as inhibitor
XX FT /note- "of HIV binding (claim 6)"
XX FT 185..209 /note- "Glu/Asp-rich region, useful as inhibitor
XX FT /note- "of HIV binding (claim 6)"

```

```

FT Region 234..271
FT /note- "Glu/Asp-rich region, useful as inhibitor
FT of HIV binding (claim 6)"
XX
XX WO9840480-A1.
XX
XX 17-SEP-1998.
XX
XX 12-MAR-1998; 98WO-EP01409.
XX
XX 12-MAR-1997; 97US-0040969.
XX
XX (CNRS ) CENT NAT RECH SCI.
XX (INSP ) INST PASTEUR.
XX
XX Briand J, Callebaut C, Guichard G, Hovanessian A;
XX PI Jacotot E, Krust B, Muller S;
XX DR WPI; 1999-034588/03.
XX DR N-PSDB; AAV71742 AND AAV71745.
XX
XX New isolated V3 loop HIV receptor - comprises p95/nucleolin,
XX P40/PHAPRI and P30/PHAPI proteins, used to develop products for the
XX treatment and prevention of HIV infection
XX
XX Disclosure; Fig 49(1); 267pp; English.
XX
XX This is the amino acid sequence of the p95 (nucleolin) protein of
XX the newly identified V3 loop HIV receptor. This novel protein
XX complex receptor for HIV retroviruses consists of an association of
XX 3 proteins, named p95/nucleolin, p40/PHAPRI and p30/PHAPI (see
XX AAM84052-54), that are implicated as cofactors in the process of
XX HIV entry into cells. Genomic and cDNA sequences for these proteins
XX are provided in AAV71742-45. The V3 loop HIV receptor proteins were
XX isolated from human CD4+ CEM T-cell extracts using an affinity
XX matrix containing either the pseudopeptide 51(KpsICH2)NPR)-template
XX assembled synthetic peptide or a synthetic V3 loop peptide (see
XX AAM84055). p95 was identified as nucleolin by amino acid sequence
XX analysis. The invention also concerns peptidic or non-peptidic
XX molecules having the ability to alter and/or prevent the binding of
XX the novel HIV receptor to the HIV retrovirus, and to pharmaceutical
XX and diagnostic compositions containing such molecules. Methods are
XX provided for screening for new active molecules, and to methods of
XX screening genetic defects in the expression of the V3 loop HIV
XX receptor in individuals that survive long-term HIV infection or who
XX are HIV-resistant. Such genetically defective polynucleotides can
XX be used in gene therapy.
XX
XX SQ Sequence 707 AA:
XX
Query Match 29.6%; Score 64; DB 20; Length 707;
Best Local Similarity 35.3%; Pred. No. 6.1;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;
QY 1 EDDERSTDSSQCCSSEDEDFEETAQVSPRGKE 34
DB 187 EDDDEDDDDDDDDDEDDSEEAETTPAKGKK 220

Search completed: June 17, 2003, 11:55:56
Job time : 12.8256 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:51:11 ; Search time 8.7157 Seconds
(Without alignments)
827.825 Million cell updates/sec

Title: US-09-991-681-29
Perfect score: 180
Sequence: 1 SFQSSSTPTGTGFGSKETPSEDROSREHMGES 35

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPREMBL_21:*
 - 2: sp_archaea:*
 - 3: sp_bacteria:*
 - 4: sp_fungi:*
 - 5: sp_human:*
 - 6: sp_invertebrate:*
 - 7: sp_mammal:*
 - 8: sp_mhc:*
 - 9: sp_organelle:*
 - 10: sp_phage:*
 - 11: sp_plant:*
 - 12: sp_prodent:*
 - 13: sp_virus:*
 - 14: sp_vertebrate:*
 - 15: sp_unclassified:*
 - 16: sp_virus:*
 - 17: sp_bacteriopl:*
 - 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	180	100.0	096CH9	096CH9 homo sapien
2	180	100.0	09ULH6	09ULH6 homo sapien
3	61	33.9	09VHT1	09VHT1 drosophila
4	59.5	33.1	09BLU9	09BLU9 leishmania
5	59	32.8	1403 5	09VPO4
6	58	32.2	252 11	09CXS4
7	57	31.7	118 3	096TV3
8	57	31.7	345 5	09VJ23
9	56	31.1	621 10	09M010
10	56	31.1	2047 4	09M010
11	55.5	30.8	468 4	096I17
12	55	30.6	89 16	097I67
13	55	30.6	656 5	096J32
14	54.5	30.3	2514 16	09JY30
15	54	30.0	322 4	075370
16	54	30.0	337 13	091708

17	54	30.0	351 16	099Z56	099Z56 streptococc
18	54	30.0	525 12	083696	083696 measles vir
19	54	30.0	600 5	09VE96	09VE96 drosophila
20	54	30.0	712 10	09M5P5	09M5P5 arabidopsis
21	54	30.0	715 10	09SYH1	09SYH1 arabidopsis
22	54	30.0	956 4	09UQ39	09UQ39 homo sapien
23	54	30.0	1262 4	09UQ40	09UQ40 homo sapien
24	54	30.0	1808 13	042142	042142 gallus gall
25	54	30.0	2296 4	09UHA8	09UHA8 homo sapien
26	54	30.0	2752 4	09UQ35	09UQ35 homo sapien
27	54	30.0	280 3	08XUJ7	08XUJ7 neurospora
28	53.5	29.7	668 6	09GLG4	09GLG4 sus scrofa
29	53.5	29.7	1209 11	09CUE8	09CUE8 mus musculu
30	53.5	29.7	1297 4	015025	015025 homo sapien
31	53	29.4	384 5	002283	002283 caenorhabdi
32	53	29.4	390 4	013249	013249 homo sapien
33	53	29.4	614 5	091956	091956 lytechinus
34	53	29.4	814 5	09XX47	09XX47 caenorhabdi
35	53	29.4	9338 11	088778	088778 rattus norv
36	52.5	29.2	169 4	096IL9	096IL9 homo sapien
37	52.5	29.2	229 10	0850B1	0850B1 oryza sativ
38	52.5	29.2	300 4	094871	094871 homo sapien
39	52.5	29.2	319 10	09C5J5	09C5J5 arabidopsis
40	52.5	29.2	363 10	09LMC6	09LMC6 arabidopsis
41	52.5	29.2	447 5	08T162	08T162 dictyosteli
42	52.5	29.2	500 11	08R364	08R364 mus musculu
43	52.5	29.2	508 5	09UAN8	09UAN8 drosophila
44	52.5	29.2	618 4	096MX3	096MX3 homo sapien
45	52.5	29.2	696 5	09VN67	09VN67 drosophila

ALIGNMENTS

RESULT 1

ID 096CH9 PRELIMINARY: PRT: 592 AA.

AC 096CH9: 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical 66.4 kDa protein (Fragment).

OC Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON;

RA Strauberg R.;

RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: BC014227; AAH14227.1; -
KW	Hypothetical protein.
FT	NON_TER
SO	SEQUENCE

Query Match: 100.0%; Score 180; DB 4; Length 592;
Best Local Similarity 100.0%; Pred. No. 5,4e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 SFQSSSTPTGTGFGSKETPSEDROSREHMGES 35
|||||
|||||

RESULT 2

ID 09ULH6 PRELIMINARY: PRT: 1807 AA.

AC 09ULH6: 096P46;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE KIA1144 protein (BIG3) (Fragment).

```

GN K1AA1244.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=1057462;
RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [2]
RP SEQUENCE OF 38-1807 FROM N.A.
RA Hong W.;
RT "K1AA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
RT subfamily of Arp GEFs.";
RL Submitted (JUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB033070; BA06558.1; -
DR EMBL: AF413080; AL04174.1; -
DR InterPro: IPR000904; Sec7.
DR SMART: SM00222; Sec7; 1.
FT NON TER
SQ SEQUENCE 1807 AA; 199921 MW; B762C29916F72CB1 CRC64;

Query Match 100.0%; Score 180; DB 4; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.8e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFQSESTPGSGFKETPSDDRSQREHMGES 35
Db 1572 SFQSESTPGSGFKETPSDDRSQREHMGES 1606

RESULT 3
O9VHY1 PRELIMINARY; PRT; 1061 AA.
AC O9VHY1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG2684 protein.
GN IDS OR CG2684.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck H., Brokstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,

```

```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003676; AF54167.1; -
DR FlyBase: FBgn002542; lds.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR000330; SNF2_N.
DR Pfam: PF00176; SNF2_N; 1.
DR Pfam: PF00176; SNF2_N; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR APF-binding; Helicase.
SQ SEQUENCE 1061 AA; 118374 MW; 536BC893B1A90509 CRC64;

Query Match 33.9%; Score 61; DB 5; Length 1061;
Best Local Similarity 51.9%; Pred. No. 6.7;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 4 SESSTPGSGFKETPSDDRSQRE 30
Db 39 SKSRPSSAGVVDETSEEESSQSE 65

RESULT 4
O9BLU9 PRELIMINARY; PRT; 1252 AA.
AC O9BLU9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Possible SNF2-related helicase.
GN L2230.06.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Zimmermann W., Wambutt R., Ivens A.C., Quail M., Rajandream M.A.,
RA Batteil B.G.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL: AL513062; CAC24685.1; -
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR000330; SNF2_N.
DR Pfam: PF00176; SNF2_N; 1.
DR Pfam: PF00176; SNF2_N; 1.
DR SMART: SM00487; DEXDC; 1.

```

DR SMART: SM00490; Helicic; 1.
 KW ARP-binding: Helicase.
 SO SEQUENCE 1252 AA; 137895 MW; 4E6D2927AB81A745 CRC64;

Query Match 33.1%; Score 59.5; DB 5; Length 1252;
 Best Local Similarity 41.0%; Pred. No. 13;
 Matches 16; Conservative 4; Mismatches 12; Indels 7; Gaps 2;

OY 4 SESSTPSTGCF---SGKETP-----SEDDRSQREHMGES 35
 Db 395 SRASTQLTGGYTGSRGKATPSTLSLSSQ000Q0RERAGSA 433

RESULT 5
 ID 09VPO4 PRELIMINARY; PRT; 1403 AA.
 AC 09VPO4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE CG4297 protein.
 GN CG4297.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Aamati P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champs C.R., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abell J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K.J., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferrar C., Ferrier S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Klumbe B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
 Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paolel J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheelert F., Shen H.,
 Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weisslock G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Glids R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003589; AAF51491.1;
 DR FlyBase: FBgn0031258; CG4297.
 SO SEQUENCE 1403 AA; 150676 MW; D7CE38B1D0C4ACDF CRC64;

Query Match 32.8%; Score 59; DB 5; Length 1403;
 Best Local Similarity 38.1%; Pred. No. 17;
 Matches 16; Conservative 5; Mismatches 13; Indels 8; Gaps 1;

OY 1 SFQSSSTPSTGSGFSGKET-----PSEDDRSQREHMG 34
 Db 1104 SSSSSSTSSLSLAGEESDENLALAEKLSQGTGHE 1145

RESULT 6
 ID 09CX54 PRELIMINARY; PRT; 252 AA.
 AC 09CX54;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 3110013H01R1k protein.
 GN 3110013H01R1k.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Glassl C., King B., Kochiya H.,
 Kuehl P., Lewis S., Matsuo Y., Mikado I., Pesole G., Quackenbush J.,
 Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
 Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohlsuki S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 DR Nature 409:685-690(2001).
 DR EMBL: AK014046; BAB29130.1;
 DR MGD: MGI:1920389; 3110013H01R1k.
 SO SEQUENCE 252 AA; 27482 MW; FDEADB1DE2A69961 CRC64;

Query Match 32.2%; Score 58; DB 11; Length 252;
 Best Local Similarity 45.8%; Pred. No. 3.6;
 Matches 11; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

OY 10 STGSGSGKETPSEDDRSQREHMG 33
 Db 20 ATGCGSGESPAQSRSTRTRAGAG 43

RESULT 7
 ID 096TV3 PRELIMINARY; PRT; 118 AA.
 AC 096TV3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical zinc finger protein (Fragment).
 GN MV129.
 OS Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Pleurotaceae; Pleurotus.
 OX NCBI_TaxID=5322;
 RN [1]

```

RP SEQUENCE FROM N.A.
RA Pisabarro A.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Park S.K.;
RT "Analysis and genetic mapping of genes expressed in the lamellae of
   the edible basidiomycete Pleurotus ostreatus."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR InterPro: IPR002328; ADH_zinc.
DR InterPro: IPR002893; ZnF_MYND.
DR Pfam: PF01753; ZF_MYND_1.
DR PROSITE: PS00059; ADH_ZINC; UNKNOWN_1.
FT NON_TER 1
FT TER 118
SQ SEQUENCE 118 AA; 13189 MW; E25005B7C57BF772 CRC64;

Query Match 31.7%; Score 57; DB 3; Length 118;
Best Local Similarity 42.4%; Pred. No. 2.2;
Matches 14; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

2 FOSSESTPGSGKGTPEDDRSQSRHMG 34
DB 44 FSESSASSISGTGVSSEDRTISGHEGE 76

RESULT 8
09VJ23 PRELIMINARY; PRT; 345 AA.
AC 09VJ23;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG15169 protein.
GN CG15169.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.E., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003661; AAF53734.1; -.
DR FlyBase: FBgn0032734; CG15169.
SQ SEQUENCE 345 AA; 37638 MW; EA17DC0CBDBEB8A1 CRC64;

Query Match 31.7%; Score 57; DB 5; Length 345;
Best Local Similarity 38.7%; Pred. No. 7.1;
Matches 12; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

5 ESSTPGSGKGTPEDDRSQSRHMGES 35
DB 123 KESTGNGGAFAPQVPTDKVDQSSGHMDTS 153

RESULT 9
09M010 PRELIMINARY; PRT; 621 AA.
AC 09M010;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 68.6 kDa protein.
GN F7A7_180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Beyan M., Pohl T., Weissenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161946; CAB82282.1; -.
DR InterPro: IPR001798; Kelch.
DR Pfam: PF01344; Kelch; 6.
DE Hypothetical protein.
SQ SEQUENCE 621 AA; 68579 MW; 259E20AAF74CE73F CRC64;

Query Match 31.1%; Score 56; DB 10; Length 621;
Best Local Similarity 38.7%; Pred. No. 19;
Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

1 SFOSESTPGSGKGTPEDDRSQSRH 31
DB 475 SVEHKSSIVAGGYDKREYLTAEFDPREH 505

RESULT 10
015019 PRELIMINARY; PRT; 2047 AA.
AC 015019;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE KIAA0301 (DJ12208.4) (KIAA0301) (Fragment).
GN KIAA0301 OR DJ12208.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```


[illegible]

```

DE DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein SP0088.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=2135709; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblum T.V., Anglucci S., Dickinson T., Hickey E.K.,
RA Holt J.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.:
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007326; AAK74275.1; -.
DR TIGR: SP0088; -.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 89 AA; 9326 MW; 3D2136BE4CC51426 CRC64;

Query Match 30.6%; Score 55; DB 16; Length 89;
Best Local Similarity 31.2%; Pred. No. 3;
Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 3 QSESSTPGSGFKETPSDDRSRGHGE 34
DB 55 ESQSATKDGQTAGKETAGEDESANQTEISQ 86
      :|:|:| | :||| || : : :
      |:|:| | :||| || : : :

RESULT 13
Q963J2 PRELIMINARY: PRT; 656 AA.
ID 0963J2 AC Q963J2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Voltage-dependent calcium channel alpha13 subunit
DE (Fragment).
DE CA-ALPHA-1T OR CG4222 OR CG15899.
GN GN
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydrioidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT Mitman S.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF371281; AAK54737.1; -.
DR FlyBase: FBgn0029846; Ca-alpha-1T.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 656 AA; 72626 MW; CEI39DB829EA1CFP CRC64;

Query Match 30.6%; Score 55; DB 5; Length 656;
Best Local Similarity 37.0%; Pred. No. 28;
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 4 SSSSTPTGSGFKETPSDDRSQSRE 30
DB 100 ASSASPASGTSGASGAPGERERDRRD 126
      :|:|:| | | | | | | | :
      :|:|:| | | | | | | | :

RESULT 14
Q9JY30
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:56:11 ; Search time 47.7804 Seconds
(without alignments)
318.982 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698

Sequence: 1 RIRAAQGVFMIDTQCSPKT.....VRQAVREMLGRVRYDIIV 518

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCUTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2698	100.0	518	4	US-09-065-383-27
2	260	99.6	49	4	US-09-065-383-31
3	216	8.0	41	4	US-09-065-383-28
4	215	8.0	40	4	US-09-065-383-30
5	180	6.7	35	4	US-09-065-383-29
6	111	4.1	474	3	US-08-729-416C-1
7	111	4.1	594	3	US-08-729-416C-7
8	110	4.1	829	1	US-07-670-611-2
9	110	4.1	829	1	US-08-220-674-2
10	110	4.1	829	1	US-08-445-186-2
11	110	4.1	829	1	US-08-446-550-2
12	110	4.1	829	2	US-08-446-550-2
13	103	3.8	1098	4	US-08-923-992A-8
14	101.5	3.8	2101	1	US-08-466-390-4
15	101.5	3.8	2101	1	US-08-470-950-4
16	101.5	3.8	2101	1	US-08-467-781-4
17	101.5	3.8	2101	2	US-08-483-924-4
18	101.5	3.8	2101	4	US-09-452-294-1
19	100.5	3.7	2101	1	US-08-195-487-4
20	100.5	3.7	2101	5	PCT-US93-06160-4
21	98	3.6	733	3	US-08-725-459B-21
22	97.5	3.5	1128	4	US-08-923-992A-6
23	95.5	3.5	1164	4	US-08-923-992A-2
24	95.5	3.5	1713	3	US-08-600-982-24
25	95.5	3.5	1713	5	PCT-US94-10261A-24
26	95.5	3.5	1805	3	US-07-853-913-2
27	94.5	3.5	907	1	US-08-990-140-4

28	94.5	3.5	907	4	US-09-546-238-4	Sequence 4, Appl1
29	94.5	3.5	940	4	US-08-810-712-7	Sequence 7, Appl1
30	94.5	3.5	1579	3	US-08-755-587-184	Sequence 184, App
31	94	3.5	397	4	US-09-006-428A-2	Sequence 2, Appl1
32	94	3.5	397	4	US-09-006-428A-19	Sequence 19, Appl1
33	94	3.5	2391	2	US-08-446-855A-2	Sequence 2, Appl1
34	94	3.5	2391	4	US-09-150-741-2	Sequence 3, Appl1
35	93.5	3.5	984	1	US-08-257-073-3	Sequence 3, Appl1
36	93.5	3.5	984	2	US-08-184-009-120	Sequence 120, App
37	93.5	3.5	984	2	US-08-458-356-120	Sequence 120, App
38	93.5	3.5	984	4	US-08-460-736-120	Sequence 120, App
39	93.5	3.5	989	4	US-08-213-419B-2	Sequence 2, Appl1
40	93.5	3.5	989	4	US-08-213-419B-4	Sequence 4, Appl1
41	93	3.4	1038	4	US-08-334-119A-2	Sequence 2, Appl1
42	92.5	3.4	793	4	US-09-588-256-10	Sequence 10, Appl1
43	92.5	3.4	906	2	US-08-609-220A-9	Sequence 9, Appl1
44	91.5	3.4	1164	4	US-08-923-992A-10	Sequence 10, Appl1
45	91.5	3.4	2842	1	US-07-741-940-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN

APPLICANT: GRANDOS, EDWARD N.

APPLICANT: HODGES, STEVEN C.

APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/065,383

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/642,385

FILING DATE: 23-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 518 amino acids

✓
no DP-
check

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 2698; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 2e-277;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRAAQOVFMIDTCSPKTPNNFHAOSCOLIIELPDEKNGHTKKSVSREIVSLL 60
DB 1 RIRAAQOVFMIDTCSPKTPNNFHAOSCOLIIELPDEKNGHTKKSVSREIVSLL 60
QY 61 SHOVLQWYDILLEEYKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDLIDSYR 120
DB 61 SHOVLQWYDILLEEYKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDLIDSYR 120
QY 121 TAREFDTSPGLKCLIKVSGIGGANLTYROSAMSNITFHALVCVLTNOETITAEQYK 180
DB 121 TAREFDTSPGLKCLIKVSGIGGANLTYROSAMSNITFHALVCVLTNOETITAEQYK 180
QY 181 VLFEDDERSTDSOQCSSEDEDIFFETAOVSPPRCKEKROWRAMPLLSVOPVSNAWYV 240
DB 181 VLFEDDERSTDSOQCSSEDEDIFFETAOVSPPRCKEKROWRAMPLLSVOPVSNAWYV 240
QY 241 LVKRLHKCMELCNNTYIOMHDLNLCMEEPLEKDPFFILPSPQSESTSTGSGSKE 300
DB 241 LVKRLHKCMELCNNTYIOMHDLNLCMEEPLEKDPFFILPSPQSESTSTGSGSKE 300
QY 301 TPSEDDRQSREHMGESLSLKAGGDLPLPSPKVEKDPKSKKEMWENAGKITMYAAD 360
DB 301 TPSEDDRQSREHMGESLSLKAGGDLPLPSPKVEKDPKSKKEMWENAGKITMYAAD 360
QY 361 KITSKLMTYKKRKQOHLNLSAPKVEKGEPLGPGQDSPLIQRPQHLMDQGMHS 420
DB 361 KITSKLMTYKKRKQOHLNLSAPKVEKGEPLGPGQDSPLIQRPQHLMDQGMHS 420
QY 421 FSAGPELLROKRPSSGSGSSLSVSRDAEAOIOAMTMVLTVINQIILPDQFTALQ 480
DB 421 FSAGPELLROKRPSSGSGSSLSVSRDAEAOIOAMTMVLTVINQIILPDQFTALQ 480
QY 481 PAVFPCISQLTCHVTDIRQAVREMLGRVGVYDIIV 518
DB 481 PAVFPCISQLTCHVTDIRQAVREMLGRVGVYDIIV 518

RESULT 2
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 9.6%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.3e-20;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 EPLGPRGQDSPLLQRPQHLMDQGMHSFSAGPELLRODKRPRSGSTGS 441
DB 1 EPLGPRGQDSPLLQRPQHLMDQGMHSFSAGPELLRODKRPRSGSTGS 49

RESULT 3
US-09-065-383-28

Sequence 28, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383

;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/842,385
;; FILING DATE: 23-APR-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Becker, Cheryl L.
;; REGISTRATION NUMBER: 35,441
;; REFERENCE/DOCKET NUMBER: 6084.US.P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 847/935-1729
;; TELEFAX: 847/938-2623
;;
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 41 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6391543e
;; US-09-065-383-28

Query Match 8.0%; Score 216; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 8e-16;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 184 EDDERSTDSQCCSSEDEDEFEETRAVSPPRGKRRAR 224
Db 1 EDDERSTDSQCCSSEDEDEFEETRAVSPPRGKRRAR 41

RESULT 4
US-09-065-383-30
Sequence 30, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

;; REFERENCE/DOCKET NUMBER: 6084.US.P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 847/935-1729
;; TELEFAX: 847/938-2623
;;
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 40 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: No. 6391543e
;; US-09-065-383-30

Query Match 8.0%; Score 215; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 9.8e-16;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 332 SPKVEKKDPSRKKEWENAGNKIYTMADKITSKLTMEYK 371
Db 1 SPKVEKKDPSRKKEWENAGNKIYTMADKITSKLTMEYK 40

RESULT 5
US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-29

Query Match
Best Local Similarity 6.7%; Score 180; DB 4; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 SFOSESSPTSTGSGKETPSEDDRSQSRHMGES 317
Db 1 SFOSESSPTSTGSGKETPSEDDRSQSRHMGES 35

RESULT 6
US-08-729-416C-1
Sequence 1, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1

Query Match
Best Local Similarity 4.1%; Score 111; DB 3; Length 474;
Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

QY 83 PGEERTIOVPE-----AKLAGFLRYI-SMONLAVIFDLILDSYRTARE 124
Db 9 PGDEPLRPRPGTGHASDQVLGPGVTYVVKYLCIEVLRSMRSLDFSTRQITREATSRV 68

QY 125 FDTSPGLAKLKK-----VSGIGGANL-YROSAMSFNIFHALYCAVITNOETTAE 176
Db 69 CEAVPGGKAGFKRRKPPSKMLSTILGKSNLOFAGMSISLTISTASLNRPPDSKOITANH 128

QY 177 OVKKVLEED-----DERSTSSQOCSSDEDLFEETAGVSPRCKEX 218
Db 129 HMRISISASGDDPTTIVAVYAKDPVNRACHILECCDGLAQDVISIGAFELRRKQY 188

QY 219 ROWRARAPLLSVQFVSNADWWVLVKRLKLCIMELCNNTYIQMHLDLENCMEERPJFKDPDF 278

Db 189 LQCPKIPALH-DRMOSLDEFPW-----TEEDDGSOHPPYNSIPS 227
QY 279 -----FILPSFOSESSPTSTGSGKETPSEDDRSQSRHMGESISL-----LKAGG 324
Db 228 KMPPGGGLDRLKRPRIAPPTAOPAGKEQYVYGR-----HLDDTGEBMQOFPPLQGS 282

QY 325 GDLLPSPKYEKKDPKSKKEMENAGNKITYTMADYTKSLMTEYKRRKQOHLASAPK 384
Db 283 SDIYSTPEGKLHVA-PTGEAPTYYNT-QQILPQAMPAAVSS--AESSPRKLDFMKPFED 338

QY 385 EVYKKEKG-----EPILPGRGODSPILQRPQ-HLMDQGMHSHSAGELLROD 431
Db 339 ALKNQPLGVPVLSKASVBCISVSPRAPDAMLELQAEYTWYQSEKSRKEEG--LLEKD 396

QY 432 ----KRRSGSTGSSLSVYADAEAOIQAMTNVLTFLNIOIILPDQFTALPAPFPCI 487
Db 397 GDFLVKRSTNTPGSGFVLGMHNGAK-----HLLVPEGTIRT-KORVVDSTI 443

QY 488 SOLTCH 493
Db 444 SHLINH 449

RESULT 7
US-08-729-416C-7
Sequence 7, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-7

Query Match
Best Local Similarity 4.1%; Score 111; DB 3; Length 594;
Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

QY 83 PGEERTIOVPE-----AKLAGFLRYI-SMONLAVIFDLILDSYRTARE 124
Db 129 PGDEPLRPRPGTGHASDQVLGPGVTYVVKYLCIEVLRSMRSLDFSTRQITREATSRV 188

```

OY 125 FDTSPGLCKLTK-----VSGIGGAANL-YROSAMSFNIYHALVCAVLNTOETIAE 176
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 169 CEAVGAGAKFRKRPPSKMLSSILGKSNLOFAGNSILITASTILNRTPOSKQIINAH 248
OY 177 QVKKVLEF-----DERSTDSQOCSEDEDFEETAQVSPRGKEX 218
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 249 HNRISIFASGDDPDTDYAVAYAKDPVNRRAHCHIECCDGLADVIGSIGAFELRFKQY 308
OY 219 RQWRARMLLSVQPSNADWMLVLRHLKLCMELCNNTYQMLDLNENCMEEPIFKDPF 278
DB 309 LQCPKIPALH-DRMQSDPEW-----TEEGGSDHPHYNSIPS 347
OY 279 -----FILSFQSESTPTSGFGSKETPSEDDRSOREHMGESLS-----LKAGG 324
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 348 KMPRGGFLDPLTKRPHAPDPAQAGKEQTYGR-----HLGTFEGDMQOTLRLGCS 402
OY 335 GDLLPSPVKYKKDPSRRKKEMENAGNKITYMAADKTI SKLMEYKKRKQOHNLSAPPK 384
DB 403 SDIYSTPBGKLVHA-PTGEAPTYVNT-QOIPQAMPAAVSS--AESPRKDLFDKMPED 458
OY 365 EVKVKKG-----EPILGRGDSPLQRPQ-HLMDQGMRSFSAGPPELLROD 431
DB 459 ALKNQPLPVLVSKAASVECTISPVSPRPAKMLELQMETWYQGEKSRKEAG--LLEKD 516
OY 432 ----RRPSGTSGLSVSRDAEAQIOAMTNMVLTVLNOIQLPDQFTTALQPAVFCI 487
DB 517 GDFLVKRSPTNGSVLTMGMHNGAK-----HLLVDEBGTIRT-KDNVFPDSI 563
OY 488 SOLTCH 493
DB 564 SHLINH 569

```

RESULT 8
US-07-670-611-2
; Sequence 2, Application US/07670611
; Patent No. 5330892
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: White, Raymond
; APPLICANT: Nakamura, Yusuke
; TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch et al.
; STREET: 1001 G Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/670,611
; FILING DATE: 19910313
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.33981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-07-670-611-2

Query Match
Best Local Similarity 18.4%; Pred No. 0.022;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

OY 81 PSPGEKTI---QVPEAKLAFRLYISMQNLVIFDILLDSYRRAREFDSPLGKCLK 136
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 307 PSTGELSTSSSSNDIPAKIAERVK-----LSKTRSSSSSDRPVLGS 349
OY 137 KVSIGGAANLYROSAMSE-----NIYFHALVCAVLNTOETIAEQVKKVLEF 185
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 350 EISSIGVSSVAEHLHSLDSCSNIOEIFOTLYSHG-----SAISESKIRE--FEV 398
OY 186 DERSTDSQOCSEDEDFEETAQVSPRGKEXRQWRARMLLSVQPSNADWMLVLR 245
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 399 ETERLNSRIEHLKSONDLITLLE-----ECKSNAERMSHLVGKYESNATALALAQ 451
OY 246 HRLCMELCNNTYQMLDLNENCMEEPIFKGDPFLLPSFQSESTPTSGFGSKETPSE 304
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 452 SEQIEAYELL-----ALAESEQSLILGQFRAVGSSPGDQSGDENITQW 498
OY 305 ----DDRQSRHMGESLSLRAG--GDLLPSPVKYKKDPSRRKKEMENAGNKITYMA 358
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 499 LKRAHDCRKTAEAKKALLMKLDGSCGGAFAVAGCSYQ-----WESLSNSHST 549
OY 359 ADKTSKLMTEYKKRKQOHNLSAPREYVVERKG-----EPILG---PRQD 401
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 550 TSSSTASSCDTEFTKEDEO-RLKDYIOQLKNDAAAKLTMLELSHIDPLSDVAVPRGDS 608
OY 402 SPL-LQRP---OHLMDQGMRSFSAGPELLRODKRPNRSGTSSLSVSRDAEAQIOAM 457
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 609 QRLDENAVLMQELMAKMEEMALKAQVLLERKK-----ALELKLSTREAQBOAY 660
OY 458 TMNVLTVLNOIQLPDQFTTALQ-----PAVPCIS---OLTCHVTDR 498
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 661 LVHIEHLSEVEQEKQRRSLSTSSGSKDPKQECADASPALSLAELRTCSBNEIA 720
OY 499 VR--QAVREMLGRVGRVDII 517
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
DB 721 AFTNAIRREKKLKAQVQELV 741

```

RESULT 9
US-08-220-674-2
; Sequence 2, Application US/08220674
; Patent No. 5571905
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: White, Raymond
; APPLICANT: Nakamura, Yusuke
; TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch et al.
; STREET: 1001 G Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/220,674
FILING DATE: 31-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/670,611
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,33981
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 829 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-220-674-2

Query Match 4.1%; Score 110; DB 1; Length 829;
Best Local Similarity 18.4%; Pred. No. 0.022;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PPSPEKTI-----QVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFDSPGLKCLK 136
DB 307 PSTGELSTSSSSNDIPAKIAERVK-----LSKTRSESSSDRPVIGS 349
QY 137 KVSIGIGCANLYROSAMSF-----NIYFHALCAVLNQETTTAQRVKVLFED 185
DB 350 EISSIGVSSVAEHLAHSLODCSNIOEIFQTLYSHG-----SAISESKIRE--FEV 398
QY 186 DERSTDSQCCSDEDEDFEETAOVSPPRGKEKRWARMPLLSVQPVSNADWVLVRL 245
DB 399 ETERLNSRIEHLKSONDILLTTL-----ECKSNARMMSLVGKESNATARLALQY 451
QY 246 HKLCMELCNNTYIOMHLDELNCEMEPPPIFKGDPFLLPSFOGESSPTSGSGKETPSE- 304
DB 452 SEQCIKAEVELL-----ALAESQSLILGFRAGVGSSPGDSDGDNITQM 498
QY 305 -----DNRSQSERHMGESLSLAKG--GDLLLPSPKYKKDPSSRKKKEWENAGNKIYMA 358
DB 499 LKRAHDCRKTENAKALMLKLDSCGGAFAVAGSVQ-----WESTLSNSHTST 549
QY 359 ADKTIKLMTEYKKRKQOHNLSAPPKEKVEKKG-----EPLG-----PRGQD 401
DB 550 TSSTASSCDTFTFEDEQ-RLKDYIQQLKNDRAAVKLTMLESHIHDPILSYDVKPRGDS 608
QY 402 SPL-LQRP--QHLMDQOMRHSFSAGPELLRODKRPRSGSTGSLSVDAEAQIQAW 457
DB 609 QRLDLENAVLMQELAMKEEMAEKALQYLLEKEK-----ALEKLTSTREAOQAY 660
QY 458 TNAVLYLNIQIQLPDQFTALQ-----PAVFPCTIS-----QLTCHVIDIR 498
DB 661 LVHIEHKSEVEDEQEMRSLSTSSGSKDKPKECADASPALISALERTTQSENELA 720
QY 499 VR--QAVREMLGRVGRVYDII 517
DB 721 AEFNNAIRREKKLARAQVELV 741

RESULT 10
US-08-445-186-2
Sequence 2, Application US/08445186
Patent No. 5576422
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.

APPLICANT: White, Raymond
APPLICANT: Nakamura, Yusuke
TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
TITLE OF INVENTION: Humans
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: Banner, Birch et al.
STREET: 1001 G Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,186
FILING DATE: 19-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,674
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: US 07/670,611
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,33981
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 829 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-445-186-2

Query Match 4.1%; Score 110; DB 1; Length 829;
Best Local Similarity 18.4%; Pred. No. 0.022;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PPSPEKTI-----QVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFDSPGLKCLK 136
DB 307 PSTGELSTSSSSNDIPAKIAERVK-----LSKTRSESSSDRPVIGS 349
QY 137 KVSIGIGCANLYROSAMSF-----NIYFHALCAVLNQETTTAQRVKVLFED 185
DB 350 EISSIGVSSVAEHLAHSLODCSNIOEIFQTLYSHG-----SAISESKIRE--FEV 398
QY 186 DERSTDSQCCSDEDEDFEETAOVSPPRGKEKRWARMPLLSVQPVSNADWVLVRL 245
DB 399 ETERLNSRIEHLKSONDILLTTL-----ECKSNARMMSLVGKESNATARLALQY 451
QY 246 HKLCMELCNNTYIOMHLDELNCEMEPPPIFKGDPFLLPSFOGESSPTSGSGKETPSE- 304
DB 452 SEQCIKAEVELL-----ALAESQSLILGFRAGVGSSPGDSDGDNITQM 498
QY 305 -----DNRSQSERHMGESLSLAKG--GDLLLPSPKYKKDPSSRKKKEWENAGNKIYMA 358
DB 499 LKRAHDCRKTENAKALMLKLDSCGGAFAVAGSVQ-----WESTLSNSHTST 549
QY 359 ADKTIKLMTEYKKRKQOHNLSAPPKEKVEKKG-----EPLG-----PRGQD 401
DB 550 TSSTASSCDTFTFEDEQ-RLKDYIQQLKNDRAAVKLTMLESHIHDPILSYDVKPRGDS 608

TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-446-350-2

Query Match 4.1%; Score 110; DB 2; Length 829;
 Best Local Similarity 18.4%; Pred. No. 0.02;
 Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PPSPEKTI-----QVPEAKLGFRLYISMQLAVIFLLDSYRTAREFTDPSGLKCLK 136
 DB 307 PSTGELSTSSSSNDIPAKIAERVK-----LSKTRSESSSDRPVLS 349
 QY 137 KVGSGIGANLYROSAMF-----NIYFHALVAVLTNQETTTADOVKKVLPE 185
 DB 350 EISSTGVSSVAEHLAHSLODCSNQELFTLISG-----SAISESKIRE--FEV 398
 QY 186 DERSTSSQCCSDEDEIFEETAOVSPPRGEKROMRAMPLLSVQPVSNADWWLVKRL 245
 DB 399 ETERLNSRIEHLKSONDLITLLE-----ECKSNABRMGMVLGKESNMATLRLALQY 451
 QY 246 HKLMELCNNTIQMHLDENCMEEPRFKGDPFILPFSQSSSTPSTGSGKTPSE- 304
 DB 452 SEQCEAVEYELL-----ALAESQSLIGQFRAAGVSGSPGDSGDEINITQM 498
 QY 305 ----DQRSQREHMGESLSLKAG--GDLLPSPFKVKDPKRKKEMWENAGNKITYMA 358
 DB 499 LKRAHDCCKTAKENAKALMLDSCGAFVAVAGSVQ-----WESLSSSHST 549
 QY 359 ADKTIKMTETKRRKQOHNLSAPPEYKVEKG-----EPLG---PRGOD 401
 DB 550 TSSTASSCDTEFTKEDEQ-RLKDYIQQLKNDRAVAKTLMLESHIHIDPLSYDVKPRGDS 608
 QY 402 SPL-LQRP---QHLMDSQOMRHSFSGPELLROKRPRTSGSGSLSVDAEAOIQAM 457
 DB 609 QRLDENAVLMQELMAKKEEMAEKQLYLEKEKK-----ALEKLSTREQOQAY 660
 QY 458 TNNVLTVLNQIQLPQDTALQ-----PAVEPCIS---OLTCHVTDIR 498
 DB 661 LVHIEHLKSEVEDEQKRMRSLSSTSGSKDKPKCECADASPALSLELRTTGCENELA 720
 QY 499 VR--QAVREMLGRVGVDDI 517
 DB 721 AETNAIRREKKLARAQELV 741

RESULT 13

US-08-923-992A-8
 Sequence 8, Application US/08923992A
 Patent No. 6280738

GENERAL INFORMATION:
 APPLICANT: Tai, Joseph Y.
 APPLICANT: Blake, Milan S.
 TITLE OF INVENTION: No. 6280738-19A FC Binding Forms of the Group B
 TITLE OF INVENTION: Streptococcal Beta Antigens
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA

COMPUTER READABLE FORM:
 ZIP: 20005
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/923,992A

FILING DATE: 05-SEP-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/024,707
 FILING DATE: 06-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 8:
 * SEQUENCE CHARACTERISTICS:
 LENGTH: 1098 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-923-992A-8

Query Match 3.8%; Score 103; DB 4; Length 1098;
 Best Local Similarity 23.7%; Pred. No. 0.19;
 Matches 103; Conservative 66; Mismatches 177; Indels 88; Gaps 22;

QY 38 PDERPNGTKRSVSRELIYVSLSHQVLLQNLVLDLLEEFYKGSPPGEKTIQVPEAKLA 97
 DB 199 PDKKEDAEVK---VREBLGRLFS--TKAGLDEIDQEHVKKTSSENTQKVE--- 247
 QY 98 GFLRYI-SMONLAVIFDLDSYRTAREFTDPSGLKCL-----LKVSGIGANLYROS 151
 DB 248 ---HANSLOMLAQKSLSELDK-ATTNQAQVKNQFLENQAKLEIPLKENTVKKLYK 303
 QY 152 AMSFNHYHALVAVLTNQETTTAEQVKVLFEDERSTD---SSQCCSDEDEIFEET 207
 DB 304 AMSESL-----EQVEKELHNSANLLEDLVAKSKETIVREYEGKLNQS 345
 QY 208 AOVSPRGEKROMRAMPLLSV-----QPVSNADWWLVKRLHKLMELCNNTIQMHLD 262
 DB 346 KNL--PELKOLE-BAHSKLQVVEDFRKKRTTSQVTPKRLKRLDLANENN--QOKTE 400
 QY 263 LENCMEPPRFKGPFFILPFSQSSSTPSTGFSG---KTPSEDOH-----SOSREH 313
 DB 401 LTVSPENTTYEGEDVKFTYAKSDKT--TLDSFDLLTKNPSVSDRISNTYKTNNDNH 458
 QY 314 MGESLSLAKGGDLLPSPVKVEKKDPSRKKEWENAGNKITYMAADKTIKLTETKKR 373
 DB 459 KIAELTIK---NLKLNESQTVTLKAK-----DDSGNV---EXTFT--ITVOKKE 500
 QY 374 KOHNLSAFPREVKYKGEPLGPRGDSPLLRQPHLMD--OGOMRHSFSGPELLROD 431
 DB 501 EKQVPKTPQKDSKTEER-VQDERKSNKNOL---QELIKSAQDELEKLEKAIKELMEQP 556
 QY 432 KRPRSGTSSSLV 445
 DB 557 EIPSNPEYGIQKSI 570

RESULT 14

US-08-466-390-4
 Sequence 4, Application US/08466390
 Patent No. 5686562

GENERAL INFORMATION:
 APPLICANT: TOURKATLY, GARY
 APPLICANT: LIDGARD, GRAHAM P
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TESTA, HORWITZ & THIBEAULT
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA

ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,390
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
APPLICATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-390-4

Query Match 3.8%; Score 101.5; DB 1; Length 2101;
Best Local Similarity 20.2%; Pred. No. 0.82;
Matches 103; Conservative 77; Mismatches 198; Indels 133; Gaps 21;

QY 50 VSFRIVSLSHOVLLONTLYDILLEEFVKGPSPEEKTIO-----VPEAKLA 97
DB 1227 ISSLEEVSIINQVLEKEGSEKELRYMAESESOKLESCACCROROPATVPELQNA 1286
QY 98 GFL-----RYSMONLAVIFDILLDSYRAREPDTSPGLCKLKKVSGIG 143
DB 1287 ALLGRRRCRAGRAREKORVASENIR--QELTQAEAEELGQE--LKAQOEK----- 1336
QY 144 AANLYROSAMSFNIYFH---ALVCAYLTNQ--ETITAEQ--VKVLFEDDERSTDSQ 194
DB 1337 ---FOKEQALSTLOLEHTSTQALVSELPKAKHLCOOLQAEQAAEKRRHELEQSKQAG 1393
QY 195 QCSSEDEDIFETNOVSPRCKEKROWRAMPILSVQVSNADWVWLKRLHKLCME--- 251
DB 1394 GLRAELLRAQRELELPLRQKVAQERTAAQ-LRAEKASYAEQSLSMKKAHGLAEENR 1452
QY 252 -----LCNNYIOMHLD--LENCMEEPPIFGDPFFILPFSQESSSTPS-----TGEF 296
DB 1453 GLGERANLGRQFLEVEILDQAREKYYOELAAVRADAEFTLAEVOREAOSTARELEVMTAKY 1512
QY 297 SGKETPSEDDRSQSRH-----MGESLSLKAGGDDLLPPSPVYEK-----DP 340
DB 1513 EGAKVKVLEERQROEEROKLTAQVEELSKLADSD---QASKVOQOKLKAVOAGGES 1568
QY 341 SRKEMWENAGNKIYTMADK-----TISKLMTEYKKRKOQH-----NLSAFP 383
DB 1569 QOEAFQAOQLNEIQAOQLSQKEQAENHYKLOMEKAKTHYDAKKQONQELQOLSLBLO 1628
QY 384 KEVK-VEKKGEPPLRGODSPFLORPOHLMDOGQMRHSFSAPELLRODKPRSGSTSS 442
DB 1629 KENNELRAEAERLG-----HELOQAGL-----KTKREAQTCRH 1661
QY 443 LSVSVRAEAOIQAMTNVLTVLNQIOLLPD 473
DB 1662 LTAQVRSLEAOV-AHADQQLRDLGKFQVATD 1691

RESULT 15
US-08-470-950-4
Sequence 4, Application US/08470950
Patent No. 5698439
GENERAL INFORMATION:
APPLICANT: TOURKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P

TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,950
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
APPLICATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-950-4

Query Match 3.8%; Score 101.5; DB 1; Length 2101;
Best Local Similarity 20.2%; Pred. No. 0.82;
Matches 103; Conservative 77; Mismatches 198; Indels 133; Gaps 21;

QY 50 VSFRIVSLSHOVLLONTLYDILLEEFVKGPSPEEKTIO-----VPEAKLA 97
DB 1227 ISSLEEVSIINQVLEKEGSEKELRYMAESESOKLESCACCROROPATVPELQNA 1286
QY 98 GFL-----RYSMONLAVIFDILLDSYRAREPDTSPGLCKLKKVSGIG 143
DB 1287 ALLGRRRCRAGRAREKORVASENIR--QELTQAEAEELGQE--LKAQOEK----- 1336
QY 144 AANLYROSAMSFNIYFH---ALVCAYLTNQ--ETITAEQ--VKVLFEDDERSTDSQ 194
DB 1337 ---FOKEQALSTLOLEHTSTQALVSELPKAKHLCOOLQAEQAAEKRRHELEQSKQAG 1393
QY 195 QCSSEDEDIFETNOVSPRCKEKROWRAMPILSVQVSNADWVWLKRLHKLCME--- 251
DB 1394 GLRAELLRAQRELELPLRQKVAQERTAAQ-LRAEKASYAEQSLSMKKAHGLAEENR 1452
QY 252 -----LCNNYIOMHLD--LENCMEEPPIFGDPFFILPFSQESSSTPS-----TGEF 296
DB 1453 GLGERANLGRQFLEVEILDQAREKYYOELAAVRADAEFTLAEVOREAOSTARELEVMTAKY 1512
QY 297 SGKETPSEDDRSQSRH-----MGESLSLKAGGDDLLPPSPVYEK-----DP 340
DB 1513 EGAKVKVLEERQROEEROKLTAQVEELSKLADSD---QASKVOQOKLKAVOAGGES 1568
QY 341 SRKEMWENAGNKIYTMADK-----TISKLMTEYKKRKOQH-----NLSAFP 383
DB 1569 QOEAFQAOQLNEIQAOQLSQKEQAENHYKLOMEKAKTHYDAKKQONQELQOLSLBLO 1628
QY 384 KEVK-VEKKGEPPLRGODSPFLORPOHLMDOGQMRHSFSAPELLRODKPRSGSTSS 442
DB 1629 KENNELRAEAERLG-----HELOQAGL-----KTKREAQTCRH 1661
QY 443 LSVSVRAEAOIQAMTNVLTVLNQIOLLPD 473
DB 1662 LTAQVRSLEAOV-AHADQQLRDLGKFQVATD 1691

Wed Jun 18 09:54:35 2003

us-09-991-681-27.ra1

Page 10

Search completed: June 17, 2003, 12:03:36
job time : 49.7804 secs

...

...

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:16 ; Search time 149,408 Seconds
(without alignments)
461.981 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRMAQGVFLDTGCSPKT.....VRAVREMLGNVGHVYDIIY 518

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2698	100.0	518	19	PS118 protein enco
2	2698	100.0	518	23	PS118 prostate mar
3	2698	100.0	1807	22	Novel human diago
4	2588	95.9	1982	22	Novel human diago
5	982	36.4	192	20	Human prostate can
6	920	34.1	180	20	Human prostate can
7	260	9.6	49	23	PS118 prostate mar
8	255	9.5	50	22	Novel human diago
9	216	8.0	41	23	PS118 prostate mar
10	215	8.0	40	23	PS118 prostate mar

11	180	6.7	35	23	AAW50811
12	177.5	6.6	2045	22	ABW61941
13	133	4.9	665	22	ABW64312
14	122	4.5	1657	22	ABW72031
15	113.5	4.2	743	22	AAE09851
16	113.5	4.2	877	22	AAE09850
17	113.5	4.2	908	22	AAE09839
18	113.5	4.2	909	22	AAE09845
19	113.5	4.2	1042	22	AAE09838
20	113.5	4.2	1043	22	AAE09844
21	111	4.1	474	18	AAW15253
22	111	4.1	594	18	AAW15256
23	111	4.1	1148	20	AAW07087
24	111	4.1	1189	22	AAU32012
25	110	4.1	792	22	AAE09853
26	110	4.1	827	22	AAE09855
27	110	4.1	829	13	AAW27549
28	110	4.1	926	22	AAE09852
29	110	4.1	957	22	AAE09841
30	110	4.1	958	22	AAE09847
31	110	4.1	961	22	AAE09854
32	110	4.1	992	22	AAE09843
33	110	4.1	993	22	AAE09849
34	110	4.1	1091	22	AAE09840
35	110	4.1	1092	22	AAE09846
36	110	4.1	1126	22	AAE09842
37	110	4.1	1127	22	AAE09848
38	108.5	4.0	872	22	AAW90753
39	108.5	4.0	1006	22	ABW63556
40	108.5	4.0	2783	22	ABW82946
41	107.5	4.0	2385	21	AAW85569
42	107.5	4.0	2608	21	AAW85574
43	107.5	4.0	2783	13	AAW23962
44	107.5	4.0	2783	13	AAW23963
45	107	4.0	750	22	AAW94129

ALIGNMENTS

RESULT 1

AAW85472

ID AAW85472 standard; Protein; 518 AA.

XX

AC AAW85472;

XX

DT 25-FEB-1999 (first entry)

XX

DE PS118 protein encoded by consensus sequence.

XX

KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.

XX

OS Homo sapiens.

XX

PN W09848054-A1.

XX

PD 29-OCT-1998.

XX

PF 23-APR-1998; 98W0-US08239.

XX

PR 23-APR-1997; 97US-0842385.

XX

PA (ABBO) ABBOT LAB.

XX

PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granados EN, Hodges SC, Kلاس MR, Krtotchiyl JD;

PI Robert-L-Strapp L, Russell JC, Stroupe SD;

XX

DR WPI; 1998-610000/51.

DR N-PSDB; AAW82812.

XX

PT New p118 nucleic acid and proteins - used for diagnosis and treatment of prostatic disease, especially cancer, and also for drug

PT screening

PS Claim 17; Page 93-94; 117pp; English.

CC The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AA92803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.

SQ Sequence 518 AA;

Query Match	100.0%	Score 2698	DB 19	Length 518
Best Local Similarity	100.0%	Pred. NO. 1e-253		
Matches 518, Conservative	0	Mismatches	0	Indels 0

QY	RIRAMAQVYMLDTQCSPTPNPNHFAOCSQJILLPDEKPNHGTKKSVEFRIVSL	60
Db	1 RIRAMAQVYMLDTQCSPTPNPNHFAOCSQJILLPDEKPNHGTKKSVEFRIVSL	60
QY	61 SHOVLONLYDILLEEYVKGPSGBEKTIOVPEAKLAFKLYKISQNLAVIFLLDLSYR	120
Db	61 SHOVLONLYDILLEEYVKGPSGBEKTIOVPEAKLAFKLYKISQNLAVIFLLDLSYR	120
QY	121 TAREPDTSPGLKCLLKKVSGIGGAANLYKQSMASNFIEFHALVCVAVLTNOETITAEQVK	180
Db	121 TAREPDTSPGLKCLLKKVSGIGGAANLYKQSMASNFIEFHALVCVAVLTNOETITAEQVK	180
QY	181 VLFEDDERSTDSSQCCSSEDEDIFEBTAVSPPRCKEKQRQWAPRMLLSVOPVSNADWV	240
Db	181 VLFEDDERSTDSSQCCSSEDEDIFEBTAVSPPRCKEKQRQWAPRMLLSVOPVSNADWV	240
QY	241 LVKRLKHLKMLCENNTIOHMLDLENCMEPPLFFKDDPFILPSFQSESPSTGGFSGKE	300
Db	241 LVKRLKHLKMLCENNTIOHMLDLENCMEPPLFFKDDPFILPSFQSESPSTGGFSGKE	300
QY	301 TPSEDDRSOSREHMGESLSKAGGDDLPPSPVYKKDPSRKKEMWENAGNKIYTMAAD	360
Db	301 TPSEDDRSOSREHMGESLSKAGGDDLPPSPVYKKDPSRKKEMWENAGNKIYTMAAD	360
QY	361 KTSISKLMTETXKKKQOONLSAPPKFVKYKKEGEPILGPGGOSPLLOPHOIMOGOMRHS	420
Db	361 KTSISKLMTETXKKKQOONLSAPPKFVKYKKEGEPILGPGGOSPLLOPHOIMOGOMRHS	420
QY	421 FSAGPELRLQDKRPRSGSTGSSLSVSVYDAEAQIOAMTNMVLTYLNOIOLIPDTFFALQ	480
Db	421 FSAGPELRLQDKRPRSGSTGSSLSVSVYDAEAQIOAMTNMVLTYLNOIOLIPDTFFALQ	480
QY	481 PAYFPCISQLTCHYVDIRVAVRQAVBMLRGVGHVYDIIV	518
Db	481 PAYFPCISQLTCHYVDIRVAVRQAVBMLRGVGHVYDIIV	518

RESULT 2
AAM50809
ID AAM50809 standard; Protein; 518 AA

AC AAM50809;

DT 01-MAY-2002 (first entry)

PS118 prostate marker partial sequence.

KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia
 KW prostatitis; human; diagnosis; therapy; vaccine.

05 Homo sapiens

PN US2001055758-A1

PD 27-DEC-2001

PF 23-APR-1998; 98US-0065383.

PR 23-APR-1997; 97US-0842385.

PA (BILL/) BILLING-MEDEL P A

PA (COPL/) COPLPITTS T L.

PA (GORD/) GORDON J.

PA (HODG/) HODGES S C.

PA (KRAT/) KRATOCHVIL J D

XX

PI Granados EN, Hódges S

XX

DR N-PSDB; ABA91651.

PT Detecting presence of

PT preventing or treating

XX
C1a1m 17: Page 42-43: Eng14ch
DS

The present sequence is that of a bi

The present sequence is that of a human prostate-specific PS18 polypeptide, as predicted from a partial consensus cDNA sequence (see ABA91651), and lacking the N-terminal region. The PS18 consensus sequence is found at least 12 times more often in prostate than in non-prostate tissue. PS18 polypeptides, including derivatives of the present sequence, polynucleotides, antibodies, agonists and inhibitors are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and treating (including by genetic immunisation), or determining the predisposition of an individual to, diseases and conditions of the prostate, such as benign prostatic hyperplasia, prostatic intraepithelial neoplasia, prostate cancer, tumours and metastases. The PS18 polypeptides can be produced by expression of PS18 polynucleotides in transfected host cells. The methods and reagents of the invention may provide an early means of detecting diseases of the prostate and may also provide new markers which can differentiate between the clinically important and unimportant prostate cancers without the use of surgery.

SQ Sequence 518 AA;

Query Match	100.0%	Score 2698	- DB 23	Length 518
Best Local Similarity	100.0%	Pred. NC 1e-253		
Matches 518	0	Mismatches 0	Indels 0	Gaps 0

QY	1	RIAMAOQVFMJLDTQCSPTPNKFNHQAOSCIIILPDEKPNGHTKSVSEFREVIVSL	60
Db	1	RIAMAOQVFMJLDTQCSPTPNKFNHQAOSCIIILPDEKPNGHTKSVSEFREVIVSL	60
QY	61	SHOVLQNLNYDILILEEFVKGSPGEEKTOVPEAKLAGFLRYISQNLAVIFDILLDSYR	120
Db	61	SHOVLQNLNYDILILEEFVKGSPGEEKTOVPEAKLAGFLRYISQNLAVIFDILLDSYR	120
QY	121	TAREFDTSPGLCKLCKLVKSGIGGANILYRQASAMSFNIYFHALVCVAVLNQETITAEQYK	180
Db	121	TAREFDTSPGLCKLCKLVKSGIGGANILYRQASAMSFNIYFHALVCVAVLNQETITAEQYK	180
QY	181	VLEFEDDERSTSSQOCSSEDEDIFEEYQAQVSPDRKEKRWRRAMPPLSVQVSNADWVY	240
Db	181	VLEFEDDERSTSSQOCSSEDEDIFEEYQAQVSPDRKEKRWRRAMPPLSVQVSNADWVY	240

```

Db      181  VLFEDDERSTDSQOCCSSEDEDFEETAQVSPPKERKQRMARPLLSVQPSNADWV 240
Oy      241  LVKRLHKLKCMELCNNTYIOMHLDLENCEMEPIFKGDPFILLPSFQSESSSTPTSGFSKE 300
Db      241  LVKRLHKLKCMELCNNTYIOMHLDLENCEMEPIFKGDPFILLPSFQSESSSTPTSGFSKE 300
Oy      301  TPSEDDRSQSRHEHMGESLSLKAGGDLPLPSPKVEKKDPKRKKEMENAGNKITYMAAD 360
Db      301  TPSEDDRSQSRHEHMGESLSLKAGGDLPLPSPKVEKKDPKRKKEMENAGNKITYMAAD 360
Oy      361  KTISKLTMEYKKRRKQOHNLAPFKVEYKKEGEPKRGDPSPLQRPQHLMDQGMHS 420
Db      361  KTISKLTMEYKKRRKQOHNLAPFKVEYKKEGEPKRGDPSPLQRPQHLMDQGMHS 420
Oy      421  FSAGPELLRQDKRRPSGSGTSSLSVSRDAEAQIOAWTMVLTVNOIOILPDQFTTALQ 480
Db      421  FSAGPELLRQDKRRPSGSGTSSLSVSRDAEAQIOAWTMVLTVNOIOILPDQFTTALQ 480
Oy      481  PAVFPCISQLTCHVTDIRVQAVREMLGKRGVYDIIV 518
Db      481  PAVFPCISQLTCHVTDIRVQAVREMLGKRGVYDIIV 518

```

RESULT 3

ABG09728 standard; Protein; 1807 AA.

ABG09728;

13-FEB-2002 (first entry)

Novel human diagnostic protein #9719.

Human: chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-0508631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dzmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS73915.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

Claim 20; SEQ ID No 40087; 103pp; English.

The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating

disorders involving aberrant protein expression or biolog
CC The polypeptide and polynucleotide sequences have applicat
CC diagnostics, forensics, gene mapping, identification of mu
CC responsible for genetic disorders or other traits to assess
CC and to produce other types of data and products dependent
CC amino acid sequences. ABG00010-ABG30377 represent novel hum
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1807 AA;

Query Match 100.0%; Score 2698; DB 22; Length 1807;

Best Local Similarity 100.0%; Pred. No. 6.9e-253;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1  RIRAAQVFMLDTCSPKTPNNFHAQSCQILTELPPEKNGHKKSVSREIYVSL 60
Db      1290 RIRAAQVFMLDTCSPKTPNNFHAQSCQILTELPPEKNGHKKSVSREIYVSL 1349
Oy      61  SHOVLQNLXDILFEFVKGPSGPEKTIQVPEAKLAGFLRYISMONLAVIFDLDSYR 120
Db      1350 SHOVLQNLXDILFEFVKGPSGPEKTIQVPEAKLAGFLRYISMONLAVIFDLDSYR 1409
Oy      121  TAREFDTSPGLKCLLKRVSIGIGAAANLYRQASAMFNITRYHALYCAVLTNOETTTAEQYK 180
Db      1410 TAREFDTSPGLKCLLKRVSIGIGAAANLYRQASAMFNITRYHALYCAVLTNOETTTAEQYK 1469
Oy      181  VLFEDDERSTDSQOCCSSEDEDFEETAQVSPPKERKQRMARPLLSVQPSNADWV 240
Db      1470 VLFEDDERSTDSQOCCSSEDEDFEETAQVSPPKERKQRMARPLLSVQPSNADWV 1529
Oy      241  LVKRLHKLKCMELCNNTYIOMHLDLENCEMEPIFKGDPFILLPSFQSESSSTPTSGFSKE 300
Db      1530 LVKRLHKLKCMELCNNTYIOMHLDLENCEMEPIFKGDPFILLPSFQSESSSTPTSGFSKE 1589
Oy      301  TPSEDDRSQSRHEHMGESLSLKAGGDLPLPSPKVEKKDPKRKKEMENAGNKITYMAAD 360
Db      1590 TPSEDDRSQSRHEHMGESLSLKAGGDLPLPSPKVEKKDPKRKKEMENAGNKITYMAAD 1649
Oy      361  KTISKLTMEYKKRRKQOHNLAPFKVEYKKEGEPKRGDPSPLQRPQHLMDQGMHS 420
Db      1650 KTISKLTMEYKKRRKQOHNLAPFKVEYKKEGEPKRGDPSPLQRPQHLMDQGMHS 1709
Oy      421  FSAGPELLRQDKRRPSGSGTSSLSVSRDAEAQIOAWTMVLTVNOIOILPDQFTTALQ 480
Db      1710 FSAGPELLRQDKRRPSGSGTSSLSVSRDAEAQIOAWTMVLTVNOIOILPDQFTTALQ 1769
Oy      481  PAVFPCISQLTCHVTDIRVQAVREMLGKRGVYDIIV 518
Db      1770 PAVFPCISQLTCHVTDIRVQAVREMLGKRGVYDIIV 1807

```

RESULT 4

ABG09731 standard; Protein; 1982 AA.

ABG09731;

13-FEB-2002 (first entry)

Novel human diagnostic protein #9722.

Human: chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS73918.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS
 PS Claim 20; SEQ ID NO 40090; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 1982 AA:
 Query Match 95.9%; Score 2588; DB 22; Length 1982;
 Best Local Similarity 99.2%; Pred. No. 4.2e-242;
 Matches 499; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 1 RIRAMQVYMLTQCSPTKPNFNDHAQSCQLITLPPDEKPNGHRTKVSFREIYVSL 60
 DB 1408 RIRAMQVYMLTQCSPTKPNFNDHAQSCQLITLPPDEKPNGHRTKVSFREIYVSL 1463
 QY 61 SHQVLLQNTYDILLEEFVKGSPGEKTIQVPAKLAGFLRYTSMONLAVIFPLLDSTR 120
 DB 1464 SHQVLLQNTYDILLEEFVKGSPGEKTIQVPAKLAGFLRYTSMONLAVIFPLLDSTR 1523
 QY 121 TAREFTSGKLKLLKRVSGIGGANLYROSAMSFNIYFHALCAVLNQTITTAQVKK 180
 DB 1524 TAREFTSGKLKLLKRVSGIGGANLYROSAMSFNIYFHALCAVLNQTITTAQVKK 1583
 QY 181 VFEDDERSTDSQCSSEDEDFEETAQVSPRGKEKQWRARMLLSVQPSNADWW 240
 DB 1584 VFEDDERSTDSQCSSEDEDFEETAQVSPRGKEKQWRARMLLSVQPSNADWW 1643
 QY 241 LVKRLHKLQMLCNNTYQMLHLENCEBPPIFGDPFLLPSFQSSSTPTSGFSKRE 300
 DB 1644 LVKRLHKLQMLCNNTYQMLHLENCEBPPIFGDPFLLPSFQSSSTPTSGFSKRE 1703
 QY 301 TSESDRSQSRHMGSLSLKAGGGLLPPSPKVKKPSRRKKEMWENAGNKIYMAAD 360
 DB 1704 TSESDRSQSRHMGSLSLKAGGGLLPPSPKVKKPSRRKKEMWENAGNKIYMAAD 1763
 QY 361 KTISKLTMEYKKRQOHNLSAFPKEVYKKGEPGPGODSPILQRPQHLMDOGMRS 420
 DB 420 KTISKLTMEYKKRQOHNLSAFPKEVYKKGEPGPGODSPILQRPQHLMDOGMRS 420

DB 1764 KTISKLTMEYKKRQOHNLSAFPKEVYKKGEPGPGODSPILQRPQHLMDOGMRS 1823
 QY 421 FSAGPELLRODKRRPSSGSSLSVSDAEAOIQAMTNMVLVNLQIQLPDQTFALQ 480
 DB 1824 FSAGPELLRODKRRPSSGSSLSVSDAEAOIQAMTNMVLVNLQIQLPDQTFALQ 1883
 QY 481 PAVFPCISQLTCHVTDIRVOAV 503
 DB 1884 PAVFPCISQLTCHVTDIRVOAV 1906
 RESULT 5
 AAY48248
 ID AAY48248 standard; Protein; 192 AA.
 XX
 AC AAY48248;
 XX
 DT 08-DEC-1999 (first entry)
 DE Human prostate cancer-associated protein 34.
 XX
 KW Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
 KW gene therapy; tissue specificity human.
 OS Homo sapiens.
 XX
 PN DE19811193-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 10-MAR-1998; 98DE-1011193.
 XX
 PR 10-MAR-1998; 98DE-1011193.
 XX
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA
 XX
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 DR WPI; 1999-519628/44.
 DR N-PSDB; AAZ33451.
 XX
 PT New nucleic acid expressed at high level in prostatic tumor tissue and
 PT encoded polypeptides, useful for treating cancer and screening for
 PT therapeutic agents.
 XX
 PS Claim 22; 128; 166bp; German.
 XX
 CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in prostatic tumor tissue and encode gene
 CC products or their fragments. The products of the invention have
 CC antitumor activity. Polypeptides (I) encoded by (A) are used: (1) for
 CC identifying agents for treatment of prostatic cancer and (11) for
 CC therapy of prostate cancer, optionally where expressed by gene therapy
 CC methods. (A) is also used to isolate full-length genes (for gene therapy)
 CC and for recombinant production of (I), which can be used to raise
 CC specific antibodies. (A) are identified by assembly of ESTs (expressed
 CC sequence tags) before they are analyzed for expression pattern (tissue
 CC specificity). This approach eliminates many of the false results, as
 CC regards tissue specificity, associated with known methods that use
 CC single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
 CC encoded by the expressed sequence tags described in the method of the
 CC invention.
 XX
 XX Sequence 192 AA:
 Query Match 36.4%; Score 982; DB 20; Length 192;
 Best Local Similarity 100.0%; Pred. No. 3.7e-87;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 MAQOVYMLTQCSPTKPNFNDHAQSCQLITLPPDEKPNGHRTKVSFREIYVSLSHOV 64
 DB 1 MAQOVYMLTQCSPTKPNFNDHAQSCQLITLPPDEKPNGHRTKVSFREIYVSLSHOV 64

OY 65 LLONLVYLLEEFVKGPSGPEKTIQVPEAKLAGFLRYISMQNLAVIFDLDSYRTARE 124
DB 61 LLONLVYLLEEFVKGPSGPEKTIQVPEAKLAGFLRYISMQNLAVIFDLDSYRTARE 120
OY 125 FDTSPGLKCLKKRVSGIGGANLYRQSA MSFNITFHALVCAVLTNOETITAEQVKVLE 184
DB 121 FDTSPGLKCLKKRVSGIGGANLYRQSA MSFNITFHALVCAVLTNOETITAEQVKVLE 180
OY 185 DDERSTSSQOC 196
DB 181 DDERSTSSQOC 192

RESULT 6
AAY48216
ID AAY48216 standard; Protein; 180 AA.

AC AAY48216;
DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 2.

KM Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
KW gene therapy; tissue specificity human.

OS Homo sapiens.

PN DE19811193-A1.

PD 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011193.

PR 10-MAR-1998; 98DE-1011193.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI: 1999-519628/44.

DR N-PSDB; AA233424.

PT New nucleic acid expressed at high level in prostatic tumor tissue and
PT encoded polypeptides; useful for treating cancer and screening for
PT therapeutic agents

PS Claim 22; 112-113; 166pp; German.

CC This invention describes novel nucleic acid sequences (A) that are
CC expressed at high level in prostatic tumor tissue and encode gene
CC products or their fragments. The products of the invention have
CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC identifying agents for treatment of prostatic cancer and (ii) for
CC therapy of prostate cancer, optionally where expressed by gene therapy
CC methods. (A) is also used to isolate full-length genes (for gene therapy)
CC and for recombinant production of (i), which can be used to raise
CC specific antibodies. (A) are identified by assembly of ESTs (expressed
CC sequence tags) before they are analyzed for expression pattern (tissue
CC specificity). This approach eliminates many of the false results, as
CC regards tissue specificity, associated with known methods that use
CC single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
CC encoded by the expressed sequence tags described in the method of the
CC invention.

XX Sequence 180 AA;

Query Match 34.1%; Score 920; DB 20; length 180;

Best Local Similarity 100.0%; Pred. No. 3.7e-81;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 MIDTQCSKPTNNPNNHQAOSCOLIILPPDEKPNHGTTRKSVSREIVSLSSHOVLQNLX 70
|||||

DB 1 MIDTQCSKPTNNPNNHQAOSCOLIILPPDEKPNHGTTRKSVSREIVSLSSHOVLQNLX 60
OY 71 DILLEEFVKGPSGPEKTIQVPEAKLAGFLRYISMQNLAVIFDLDSYRTAREPDTSPG 130
DB 61 DILLEEFVKGPSGPEKTIQVPEAKLAGFLRYISMQNLAVIFDLDSYRTAREPDTSPG 120
OY 131 LKCLLKRVSGIGGANLYRQSA MSFNITFHALVCAVLTNOETITAEQVKVLEFDDERST 190
DB 121 LKCLLKRVSGIGGANLYRQSA MSFNITFHALVCAVLTNOETITAEQVKVLEFDDERST 180

RESULT 7
AAM50813
ID AAM50813 standard; Protein; 49 AA.

AC AAM50813;
DT 01-MAY-2002 (first entry)

DE PS118 prostate marker immunogenic polypeptide.

KM PS118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostaticitis; human; diagnosis; therapy; vaccine; immunogen.

OS Homo sapiens.

PN US2001055758-A1.

PD 27-DEC-2001.

PF 23-APR-1998; 98US-0065383.

PR 23-APR-1997; 97US-0842385.

PA (BILL-) BILLING-MEDEL P A.

PA (COHE/) COHEN M.

PA (COPL/) COPLITTS T L.

PA (FRIE/) FRIEDMAN P N.

PA (GORD/) GORDON J.

PA (GRAN/) GRANADOS E N.

PA (HODG/) HODGES S C.

PA (KLAS/) KLAS M R.

PA (KRAT/) KRATOCHVIL J D.

PA (ROBE/) ROBERTS-RAPP L.

PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;

PI Russell JC, Stroupe SD;

DR WPI: 2002-187683/24.

PT Detecting presence of target PS118 polynucleotide in test sample,

PT useful for detecting, diagnosing, staging, monitoring, prognosticating,

PT preventing or treating or determining predisposition to prostate

PT disease -

PS Claim 17; Page 44; 57pp; English.

XX The present sequence is that of an immunogenic polypeptide

XX comprising amino acids 393-441 of human prostate-specific PS118

XX polypeptide (see AAM50809). A PS118 consensus sequence (see

XX ABA91651) is found at least 12 times more often in prostate than

XX in non-prostate tissue. PS118 polypeptides, polynucleotides,

XX antibodies, agonists and inhibitors are useful for detecting,

XX diagnosing, staging, monitoring, prognosticating, preventing and

XX treating (including by genetic immunisation), or determining the

XX predisposition of an individual to, diseases and conditions of the

XX prostate, such as benign prostatic hyperplasia, prostaticitis,

XX prostatic intraepithelial neoplasia, prostate cancer, tumours and

XX metastases. The PS118 polypeptides can be produced by expression

XX of PS118 polynucleotides in transfected host cells, and

XX immunogenic peptides are useful for raising PS118-specific

antibodies of diagnostic use. The methods and reagents of the invention may provide an early means of detecting diseases of the prostate and may also provide new markers which can differentiate between the clinically important and unimportant prostate cancers without the use of surgery.

Sequence 49 AA:

Query Match 9.6%; Score 260; DB 23; Length 49;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

393 EPICPRQDPSFLQRPQIMDQGMHSFSAGPELLRQDKRRSGSTGS 441

1 EPICPRQDPSFLQRPQIMDQGMHSFSAGPELLRQDKRRSGSTGS 49

RESULT 8

ABG09729 standard; Protein: 50 AA.

ABG09729;

13-FEB-2002 (first entry)

Novel human diagnostic protein #9720.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

33-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dermanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS73916.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 40088; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 50 AA:

Query Match 9.5%; Score 255; DB 22; Length 50;
Best Local Similarity 98.0%; Pred. No. 3.3e-17;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

460 MVLTVLNOIOLTPOTFTALPAVFPCTISOLTCVHTDIRVQAAREMIGR 509

1 MVLTVLNOIOLTPOTFTALPAVFPCTISOLTCVHTDIRVQAAREMIGR 50

RESULT 9

AAM50810 standard; Protein: 41 AA.

AAM50810;

01-MAY-2002 (first entry)

PS118 prostate marker immunogenic polypeptide.

PS118; prostate; marker; prostate cancer; tumour; metastasis; benign prostatic hyperplasia; prostatic intraepithelial neoplasia; prostatitis; human; diagnosis; therapy; vaccine; immunogen.

Homo sapiens.

US2001057578-A1.

27-DEC-2001.

23-APR-1998; 98US-0065383.

23-APR-1997; 97US-0842385.

(BILL/) BILLING-MEDEL P A.

(COHE/) COHEN M.

(COPL/) COPLITTIS T L.

(FRIE/) FRIEDMAN P N.

(GORD/) GORDON J.

(GRAN/) GRANADOS E N.

(HODG/) HODGES S C.

(KLAS/) KLAAS M R.

(KRAT/) KRATOCHVIL J D.

(ROBE/) ROBERTS-RAPP L.

Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J,

Granados EN, Hodges SC, Klaas MK, Kratochvil JD, Roberts-Rapp L,

Russell JC, Stroupe SD;

WPI; 2002-187683/24.

Detecting presence of target PS118 polynucleotide in test sample,

useful for detecting, diagnosing, staging, monitoring, prognosticating,

preventing or treating or determining predisposition to prostate disease.

Claim 17; Page 43; 57pp; English.

The present sequence is that of an immunogenic polypeptide comprising amino acids 184-224 of human prostate-specific PS118 polypeptide (see AAM50809). A PS118 consensus sequence (see ABG91651) is found at least 12 times more often in prostate than in non-prostate tissue. PS118 polypeptides, polynucleotides, antibodies, agonists and inhibitors are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and treating (including by genetic immunisation), or determining the predisposition of an individual to, diseases and conditions of the prostate, such as benign prostatic hyperplasia, prostatitis,

CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells, and
CC immunogenic peptides are useful for raising PS118-specific
CC antibodies of diagnostic use. The methods and reagents of the
CC invention may provide an early means of detecting diseases of the
CC prostate and may also provide new markers which can differentiate
CC between the clinically important and unimportant prostate cancers
CC without the use of surgery.

XX Sequence 41 AA: `

SO Query Match 8.0%; Score 216; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 184 EDDERSTDSQCCSEDEDFEETAOVSPPRGKRRAR 224
Db 1 EDDERSTDSQCCSEDEDFEETAOVSPPRGKRRAR 41

RESULT 10
AAM50812
ID AAM50812 standard; Protein: 40 AA.
AC AAM50812;
XX
XX 01-MAY-2002 (first entry)
DE PS118 prostate marker immunogenic polypeptide.
XX
XX PS118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
XX
XX Homo sapiens.
OS
XX US2001055758-A1.
XX
XX 27-DEC-2001.
PD
XX 23-APR-1998; 98US-0065383.
XX
XX 23-APR-1997; 97US-0842385.
PR
XX
XX (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COPL/) COPLITTIS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
XX Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD.
XX
XX WPI: 2002-187683/24.
DR
XX
XX Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX
XX Claim 17; Page 44; 57pp; English.
XX
XX The present sequence is that of an immunogenic polypeptide
CC comprising amino acids 332-371 of human prostate-specific PS118
CC polypeptide (see AAM50809). A PS118 consensus sequence (see
CC ABA91651) is found at least 12 times more often in prostate than

CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells, and
CC immunogenic peptides are useful for raising PS118-specific
CC antibodies of diagnostic use. The methods and reagents of the
CC invention may provide an early means of detecting diseases of the
CC prostate and may also provide new markers which can differentiate
CC between the clinically important and unimportant prostate cancers
CC without the use of surgery.

XX Sequence 40 AA:

SO Query Match 8.0%; Score 215; DB 23; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 SPKVEKKDPSRKKEWENAGNKITYTMAADKTIKLTMYEK 371
Db 1 SPKVEKKDPSRKKEWENAGNKITYTMAADKTIKLTMYEK 40

RESULT 11
AAM50811
ID AAM50811 standard; Protein: 35 AA.
AC AAM50811;
XX
XX 01-MAY-2002 (first entry)
DE PS118 prostate marker immunogenic polypeptide.
XX
XX PS118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
XX
XX Homo sapiens.
OS
XX US2001055758-A1.
XX
XX 27-DEC-2001.
PD
XX 23-APR-1998; 98US-0065383.
XX
XX 23-APR-1997; 97US-0842385.
PR
XX
XX (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COPL/) COPLITTIS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
XX Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD.
XX
XX WPI: 2002-187683/24.
DR
XX
XX Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX
XX

PS Claim 17; Page 44; 57pp; English.
XX
CC The present sequence is that of an immunogenic polypeptide
CC comprising amino acids 283-317 of human prostate-specific PS118
CC polypeptide (see AAM50809). A PS118 consensus sequence (see
CC ABA91651) is found at least 12 times more often in prostate than
CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells, and
CC immunogenic peptides are useful for raising PS118-specific
CC antibodies of diagnostic use. The methods and reagents of the
CC invention may provide an early means of detecting diseases of the
CC prostate and may also provide new markers which can differentiate
CC between the clinically important and unimportant prostate cancers
CC without the use of surgery.
XX

SO Sequence 35 AA;

Query Match 6.7%; Score 180; DB 23; Length 35;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 283 SFOSESTPTGTGFGSGKETPESEDDRSQSRHNGES 317
DB 1 SFOSESTPTGTGFGSGKETPESEDDRSQSRHNGES 35

RESULT 12
ABB61941
ID ABB61941 standard; Protein: 2045 AA.

XX ABB61941;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 12615.

XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, LI PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB: ABL06044.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure: SEQ ID NO 12615; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB057737-AB072072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX

SO Sequence 2045 AA;

Query Match 6.6%; Score 177.5; DB 22; Length 2045;
Best Local Similarity 29.9%; Pred. No. 3.6e-07;
Matches 60; Conservative 25; Mismatches 53; Indels 63; Gaps 7;

OY 1 RIRAAQOVEMLDYQCSF-----KTPNFDHAGSCQILIELPDEKP-----NGHT 46
DB 1639 RIVLAQOVFLSDNQREGQNAQTPS-----ASQCKL-----SDDRSYFLYPLNNGFN 1689
OY 47 KK-----SVSFREIVSLSHQVLYLQNYDILIEF-----VKGPS 82
DB 1690 SNLDNFVIRIPFKMLVVGILLANQMLQLVAKLLRLKCVQAVSTCFDNYAASAPS 1749

OY 83 PGEKTIQVPEAKLAGFLRYISMQNLAVIFPDLDSYRTAREPTSPGICLLKKVSGIG 142
DB 1750 HDYDLDERSKTEILRCYKQYL-----MSALEPDSRGLKFLMKQVSNIE 1793

OY 143 GAANLYRQASAMSPNIYFHALV 163
DB 1794 YAANLYRQMTSSMMIYIALV 1814

RESULT 13
ABB64312
ID ABB64312 standard; Protein: 665 AA.

XX ABB64312;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 19728.

XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, LI PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB: ABL08415.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure: SEQ ID NO 19728; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:48:41 ; Search time 2.10102 Seconds
(without alignments)
690.935 Million cell updates/sec

Title: US-09-991-681-29
Perfect score: 180
Sequence: 1 SFQSSSTPSTGSGKTPEDRSQSRHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	33.9	1061	LDS_DROME	P34739 drosophila
2	57.5	31.9	592	ODP2_DICDI	P36413 dictyostell
3	55.5	30.8	490	CHE1_HUMAN	000409 homo sapien
4	55	30.6	585	YMO4_SCHPO	09hgm7 schizosacch
5	54.5	30.3	568	DONS_DROME	09vna8 drosophila
6	54.5	30.3	573	EYAA_HUMAN	099503 homo sapien
7	53.5	29.7	737	KPCE_MOUSE	P16054 mus musculu
8	53.5	29.7	737	KPCE_RAT	P10466 ratulus norv
9	52.5	29.2	508	NO60_DROME	P04081 drosophila
10	52	28.9	675	ZG20_XENLA	P18714 xenopus lae
11	52	28.9	902	UBPF_HUMAN	09y468 homo sapien
12	52	28.9	977	KFMS_MOUSE	P09581 mus musculu
13	52	28.9	978	KFMS_RAT	000495 ratulus norv
14	51.5	28.6	427	CREA_ASPNG	005620 aspergillus
15	51	28.3	83	NXSI_NAJAT	P01430 naja atra
16	51	28.3	131	TRK5_ECOLI	P17908 escherichia
17	51	28.3	613	DNAR_THEVO	097b98 thermoplasma
18	51	28.3	757	TAU_HUMAN	P10636 homo sapien
19	51	28.3	1709	CHDI_HUMAN	P10466 homo sapien
20	51	28.3	1711	CHDI_MOUSE	P40201 mus musculu
21	50.5	28.1	522	NU62_HUMAN	P37188 homo sapien
22	50.5	28.1	969	ARVC_MOUSE	P98203 mus musculu
23	50	27.8	309	MAG1_HUMAN	P43355 homo sapien
24	50	27.8	671	CHEA_THEMA	056310 thermotoga
25	50	27.8	686	CHEA_RHOSH	053135 rhodobacter
26	50	27.8	1462	NKCR_HUMAN	P30414 homo sapien
27	50	27.8	1770	PMPC_CHITR	084419 chlamydia t
28	50	27.8	1912	VITI_CHICK	P87488 gallus galli
29	49.5	27.5	168	VLPD_MYCHR	049536 mycoplasma
30	49.5	27.5	515	STEF_YEAST	P06784 saccharomyc
31	49.5	27.5	736	KPCE_RABIT	P10830 oryctolagus
32	49.5	27.5	737	KPCE_HUMAN	002156 homo sapien
33	49.5	27.5	3178	YS89_CAEEL	009624 caenorhabdi

34	49	27.2	278	MA32_MOUSE	035658 mus musculu
35	49	27.2	278	MA33_RAT	035796 ratulus norv
36	49	27.2	326	1265_MOUSE	091020 mus musculu
37	49	27.2	332	1265_RAT	035986 ratulus norv
38	49	27.2	821	GYRA_BACSU	P05653 bacillus su
39	49	27.2	1127	Y855_TREPA	083827 treponema p
40	48.5	26.9	189	BCP_PEA	041001 plisum sativ
41	48.5	26.9	511	NF60_LOLPE	001240 loligo peal
42	48.5	26.9	568	1 AF9_HUMAN	P42568 homo sapien
43	48.5	26.9	656	UL25_HCMVA	P16761 human cytom
44	48.5	26.9	736	PRX1_HUMAN	092786 homo sapien
45	48.5	26.9	737	PRX1_MOUSE	P48437 mus musculu

ALIGNMENTS

RESULT 1
LDS_DROME STANDARD: PRT: 1061 AA.
ID LDS_DROME
AC P34739;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE Probable hellicase Iodestar.
GN LDS.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92009170; PubMed=1916263;
RA Girdham C.G., Glover D.M.;
RT "Chromosome tangling and breakage at anaphase result from mutations in Iodestar, a Drosophila gene encoding a putative nucleoside triphosphate-binding protein."
RT, Genes Dev. 5:1786-1799(1991).
RN [2]
RP CONCEPTUAL TRANSLATION.
RX MEDLINE=93181281; PubMed=8382805;
RA Bork P., Koonin E.V.;
RT "An expanding family of hellicases within the 'DEAD/H' superfamily."
RT Nucleic Acids Res. 21:751-752(1993).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO BE INTRODUCED IN POSITION 946 TO PRODUCE THIS ORF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X62629; CAA4496.1; ALU_FRAME.
CC PTR: A40580; A40580.
CC FLYBase: F8gn002342; lds.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR001650; Helicase_C.
CC InterPro: IPR000330; SNF2_N.
CC Pfam: PF00176; SNF2_N: 1.
CC Pfam: PF00271; helicase_C: 1.
CC SMART: SM00487; DEXDC: 1.
CC SMART: SM00490; HELIC_C: 1.
CC Nuclear protein; Helicase; ATP-binding.
CC NP_BIND 465 472 ATP (POTENTIAL).
CC SITE 603 606 DEAD BOX.
CC SEQUENCE 1061 AA; 118189 MW; E82AA64254342B80 CRC64;

Best Local Similarity 39.5%; Pred. No. 6.7;
Matches 15; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

OY 5 ESSTPSTGSGFKETPSS---EDDRSOS---REHMGES 35
DB 375 EGSSESGSFRESHSPSTEDDRKRSKREPKDSLDS 412

RESULT 4
YMO4_SCHPO STANDARD: PRT: 585 AA.
ID YMO4_SCHPO
AC 09HGM7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C543.04 in chromosome II.
GN SPBC543.04.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Galliliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Collins K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Cocks M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., O'Neill C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skelton J., Simmonds M., Squares R., Squares S., Sharp S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Meller-ner S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revela J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0171 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL391604; CAC05246.1;
CC InterPro: IPR005365; UPF0171.
CC Pfam: PF03666; UPF0171; 1.
CC DR Hypothetical protein.
CC KW SEQUENCE 585 AA; 66868 MW; 92F53DBF5AC1BE59 CRC64;
CC

Query Match 30.6%; Score 55; DB 1; Length 585;
Best Local Similarity 46.4%; Pred. No. 9.6;
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
OY 5 ESSTPSTGSGFKETPSSDDRSOSREHM 32

DB 143 ESSHPSTGSGFEVKKSSSKRSRSMLEFHV 170

RESULT 5
DONS_DROME STANDARD: PRT: 568 AA.
ID DONS_DROME
AC 09VNB8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Downstream of son gene protein homolog.
GN CG2669.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch L., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson B., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklow G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE DONSON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE003603; AAF52038.1;
CC FlyBase: FBgn0037316; CG2669.
CC DR Hypothetical protein.
CC KW

SO SEQUENCE 568 AA: 63437 MW: A7E29D4CQAFD14B0 CRC64;
 Query Match 30.3%; Score 54.5; DB 1; Length 568;
 Best Local Similarity 36.8%; Pred. No. 11;
 Matches 14; Conservative 5; Mismatches 14; Indels 5; Gaps 1;

OY 1 SFQSESTPTGCGFGKEMP-----SEDDRSQSRHMG 33
 Db 325 SFNESTTSLGPEAGEDAPPAQEDDEDDMLSLG 362

RESULT 6
 EYA3_HUMAN STANDARD: PRT: 573 AA.
 ID EYA3_HUMAN 09504: 095463;
 AC 09504: 095463; (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eyes absent homolog 3.
 GN EYA3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Embryo;
 RA MEDLINE-97172972; PubMed-9020840;
 RA Abdelhak S., Kalatzis V., Heilig R., Compain S., Samson D.,
 RA Vincent C., Weil D., Craud C., Sahly I., Lebovici M.,
 RA Bitter-Glindicz M., Francis M., Lacombe D., Vigneron J.,
 RA Charachon R., Boyen K., Bededer P., van Regemorter N.,
 RA Weissenbach J., Petit C.;
 RT "A human homologue of the Drosophila eyes absent gene underlies
 RT branchio-oto-renal (BOR) syndrome and identifies a novel gene
 RT family.";
 RT Nat. Genet. 15:157-164(1997).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Retina;
 RA MEDLINE-97202104; PubMed-9049631;
 RA Zimmerman J.E., Bul O.T., Steingrimsen E., Nagle D.L., Fu W.,
 RA Genin A., Spliner N.B., Copeland N.G., Jenkins N.A., Bucan M.,
 RA Bonini N.M.;
 RT "Cloning and characterization of two vertebrate homologs of the
 RT Drosophila eyes absent gene.";
 RT Genome Res. 7:128-141(1997).
 RL [3]
 RP SEQUENCE OF 413-531 FROM N.A.
 RA MEDLINE-99105912; PubMed-9887327;
 RA Borani G., Degrandi A., Ballabio A., Bulfone A., Bernard L.,
 RA Banti S., Gattuso C., Mariani M., Dixon M., Donati D., Metcalfe K.,
 RA Winter R., Robertson M., Axton R., Brown A., van Heyningen V.,
 RA Hanson I.;
 RT "EYA4, a novel vertebrate gene related to Drosophila eyes absent.";
 RT Hum. Mol. Genet. 8:11-23(1999).
 RL [1]
 RP FUNCTION: MAY BE INVOLVED IN DEVELOPMENT OF THE EYE.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; SEEM TO BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE EYA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y10262; CAA71311.1;
 CC EMBL: U81602; AAB42066.1;
 CC EMBL: AJ007591; CAA07814.1;
 CC Gene; HGNC:3521; EYA3.
 DR

DR MIM: 601655;
 DR InterPro: IPR001454; Hlgase/hydrlase.
 DR Pfam: PF00702; Hydrolase; 1.
 DR Developmental protein: Multigene family: Alternative splicing.
 KW VARSPLIC 126 MISSING (IN ISOFORM 2).
 FT CONFLICT 142 142 H->R (IN REF. 2).
 FT CONFLICT 151 151 L->V (IN REF. 2).
 FT CONFLICT 151 151 K->V (IN REF. 2).
 FT CONFLICT 253 253 T->S (IN REF. 2).
 FT CONFLICT 268 268 T->S (IN REF. 2).
 FT CONFLICT 283 283 K->R (IN REF. 2).
 FT CONFLICT 290 290 K->R (IN REF. 2).
 FT CONFLICT 305 305 V->L (IN REF. 2).
 FT CONFLICT 305 305 K->E (IN REF. 2).
 FT CONFLICT 345 345 K->S (IN REF. 2).
 FT CONFLICT 357 357 S->F (IN REF. 2).
 FT CONFLICT 361 361 K->E (IN REF. 2).
 FT CONFLICT 373 373 PN->SD (IN REF. 2).
 FT CONFLICT 376 376 K->N (IN REF. 2).
 FT CONFLICT 379 379 N->D (IN REF. 2).
 FT CONFLICT 381 381 N->S (IN REF. 2).
 FT CONFLICT 388 388 N->D (IN REF. 2).
 FT CONFLICT 447 447 R->K (IN REF. 1).
 FT CONFLICT 449 449 R->K (IN REF. 1).
 FT CONFLICT 457 457 D->N (IN REF. 1).
 FT CONFLICT 473 473 R->K (IN REF. 1).
 FT CONFLICT 480 480 L->P (IN REF. 1).
 FT CONFLICT 487 487 V->L (IN REF. 1).
 FT CONFLICT 500 500 E->K (IN REF. 1).
 FT CONFLICT 524 524 SRF->TSL (IN REF. 1).
 FT CONFLICT 530 530 V->L (IN REF. 1).
 SO SEQUENCE 573 AA: 62558 MW: EBE1F6435295CACS CRC64;

Query Match 30.3%; Score 54.5; DB 1; Length 573;
 Best Local Similarity 31.9%; Pred. No. 11;
 Matches 15; Conservative 6; Mismatches 11; Indels 15; Gaps 1;

OY 1 SFQSE-----SSTPTGCGFGKEMPSEDDRSQSRHMG 32
 Db 236 TYQSEKPSYMAPAPAAQKLSGDPSTPSLSQTPSKOTDDQSRNM 282

RESULT 7
 KPCF_MOUSE STANDARD: PRT: 737 AA.
 ID KPCF_MOUSE 016054;
 AC 016054;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein kinase C, epsilon type (EC 2.7.1.-) (nPKC-epsilon).
 GN PKCE OR PKCE OR PKCEA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-89137541; PubMed-2917656;
 RA Schapp D., Parker P.T., Bristol A., Kriz R., Knopf J.;
 RT "Unique substrate specificity and regulatory properties of
 RT PKC-epsilon: a rationale for diversity.";
 RT FEBS Lett. 243:351-357(1989).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE-98127436; PubMed-9467942;
 RA Wang Q.J., Acs P., Goodnight J., Blumberg P.M., Mischak H.,
 RA Mushinski J.F.;
 RT "The catalytic domain of PKC-epsilon, in reciprocal PKC-delta and
 RT epsilon chimeras, is responsible for conferring tumorigenicity to
 RT NIH3T3 cells, whereas both regulatory and catalytic domains of
 RT PKC-epsilon contribute to in vitro transformation.";
 RT Oncogene 16:53-60(1998).
 RL [3]
 RP SEQUENCE FROM N.A.
 DR

CC TISSUE-Brain:
 RA Wheeler D.L.:
 CC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,
 CC SERINE- AND THREONINE-SPECIFIC ENZYME.
 CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
 CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
 CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF028009; AAB84189.1; -
 DR EMBL; AF325507; AAG53692.1; -
 DR PIR; S02270; KIMSCE.
 DR HSSP; P28867; IPTQ.
 DR MGD; MGI:97599; PKce.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase_1.
 DR Pfam; PF00130; DAG_PE-bind_2.
 DR Pfam; PF00168; C2_1.
 DR Pfam; PF00433; Pkinase_C_1.
 DR PRINTS; PR00008; DAGPEDOMAIN.
 DR PRODOM; PD000001; Euk_pkinase_1.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00133; S_TK_X_1.
 DR SMART; SM00220; S_TKc_1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Repeat; ATP-binding; Transferase; Phosphorylation;
 KW Serine/threonine-protein kinase; Phorbol-ester binding; Zinc.
 FT DOMAIN 1 99
 FT DOMAIN 170 220 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 2.
 FT DOMAIN 408 668 PROTEIN KINASE.
 FT NP_BIND 414 422 ATP (BY SIMILARITY).
 FT BINDING 437 437 ATP (BY SIMILARITY).
 FT ACT_SITE 532 532 BY SIMILARITY.
 FT MOD_RES 703 703 PHOSPHORYLATION (AURO-) (POTENTIAL).
 FT MOD_RES 710 710 PHOSPHORYLATION (AURO-) (POTENTIAL).
 FT SEQUENCE 737 AA: 83560 MW: 74888 Cys99F57 Cys64;
 Query Match 29.7%; Score 53.5; DB 1; Length 737;
 Best Local Similarity 48.0%; Pred. No. 20;
 Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;
 Qy 4 SESSTPTGSGSGKETSSEDRSQS 28
 Db 327 AESPPASG-----NSPEEDRSKS 346
 RESULT 8
 KPE_RAT
 ID KPE_RAT STANDARD: PRT; 737 AA.

AC P09216;
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 13-JUL-1999 (Rel. 38, Last annotation update)
 DE Protein kinase C, epsilon type (EC 2.7.1.37) (nPKC-epsilon).
 GN PRKCE OR PKCE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=88198270; PubMed=2834397;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igarashi K., Nishizuka Y.;
 RT "The structure, expression, and properties of additional members of
 RT the protein kinase C family.";
 RL J. Biol. Chem. 263:6927-6932(1988).
 RN [2]
 RP SEQUENCE OF 135-297 FROM N.A.
 RX MEDLINE=88083621; PubMed=3691811;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igarashi K., Nishizuka Y.;
 RT "Identification of three additional members of rat protein kinase C
 RT family: delta-, epsilon- and zeta-subspecies.";
 RL FEBS Lett. 226:125-128(1987).
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,
 CC SERINE- AND THREONINE-SPECIFIC ENZYME.
 CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
 CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
 CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M18331; AAA41872.1; -
 DR PIR; B28163; KIRTCe.
 DR HSSP; P28867; IPTQ.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase_1.
 DR Pfam; PF00130; DAG_PE-bind_2.
 DR Pfam; PF00168; C2_1.
 DR Pfam; PF00433; Pkinase_C_1.
 DR PRINTS; PR00008; DAGPEDOMAIN.
 DR PRODOM; PD000001; Euk_pkinase_1.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00133; S_TK_X_1.
 DR SMART; SM00220; S_TKc_1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Repeat; ATP-binding; Transferase; Phosphorylation;
 KW Serine/threonine-protein kinase; Phorbol-ester binding; Zinc.
 FT DOMAIN 1 99
 FT DOMAIN 170 220 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 2.

Query Match	Best Local Similarity	Score	DB 1:	Length	737:
Matches 12: Conservative	5: Mismatches	3: Indels	5: Gaps		
<p>4 SESSTPGTGCGRKTEPSEDORS 28</p> <p>11 111111111</p> <p>327 AESTQPASG-----NSPSEDRSKS 346</p>					
<p>RESULT 9</p> <p>NO60_DROME STANDARD: PRT: 508 AA.</p>					
ID NO60_DROME	STANDARD:	PRT:	508 AA.		
AC 044081: G9V325;					
DT 15-DEC-1998 (Rel. 37, Created)					
DT 15-DEC-1998 (Rel. 37, Last sequence update)					
DT 15-JUN-2002 (Rel. 41, Last annotation update)					
DE Nucleolar protein Ar band 60B (Minify protein).					
GN NOP60B OR MFL OR CG3333.					
OS Drosophila melanogaster (Fruit fly):					
OC Eukaryota, Metazoa, Arthropoda, Mandibulata, Pancrustacea, Hexapoda:					
OC Insecta, Pterygota, Arthropoda, Endopterygota, Diptera, Brachycera:					
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
OX NCBI_TaxID:7227;					
11					
RP SEQUENCE FROM N.A.					
RX MEDLINE=99045289; PubMed=9829924;					
RA Phillips B., Billin A.N., Cadwell C., Buchholz R., Erickson C.,					
RA Merriam J.R., Carbon J., Poole S.J.;					
RT "The NOP60B gene of Drosophila encodes an essential nucleolar protein					
RT that functions in yeast."					
RL Mol. Gen. Genet. 260:20-29(1998).					
12					
RP SEQUENCE FROM N.A.					
RX MEDLINE=99189256; PubMed=10087258;					
RA Giordano E., Peluso I., Senger S., Furia M.;					
RT "Minify, a Drosophila gene required for ribosome biogenesis."					
RT J. Cell Biol. 144:1123-1133(1999).					
13					
RP SEQUENCE FROM N.A.					
AC STRAIN-Berkeley;					
RX MEDLINE=20196006; PubMed=10731132;					
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Goebye J.D.,					
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,					
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,					
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,					
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,					
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,					
RA Abrell J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,					
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,					
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,					
RA Botkova D., Botchan M.R., Bouck J., Broxstein P., Brotlier P.,					
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,					
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,					
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,					
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,					
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,					
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,					
RA Harris N.L., Harey D., Helman T.J., Hernandez J.R., Houck J.,					
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,					
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,					
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,					
RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,					
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,					
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,					

RA Mount S.R., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Port V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sliden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spryler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Sylvester R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarnon D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Wotley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveril J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Glabbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2165-2195(2000).
CC -I- FUNCTION: PLAYS A CENTRAL ROLE IN RIBOSOMAL RNA PROCESSING.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF017230; AAC97117.1; -
DR EMBL; AF089837; AAD16092.1; -
DR EMBL; AE003463; AAF47178.1; -
DR FlyBase; FBgn0023184; Nop60B.
DR InterPro; IPRO04802; CBf5.
DR InterPro; IPRO02478; PUA.
DR InterPro; IPRO02501; Trub_N.
DR Pfam; PF01472; PUA; 1.
DR Pfam; PF01509; Trub_N; 1.
DR SMART; SM00359; PUA; 1.
DR TIGRFAMs; TIGR00425; CBf5; 1.
KW Nuclear protein; tRNA processing.
FT DOMAIN 8 16 POLY-LYS.
FT DOMAIN 450 453 POLY-ALA.
FT DOMAIN 461 471 POLY-LYS.
FT DOMAIN 489 499 POLY-LYS.
SQ SEQUENCE 508 AA; 56830 MW; 3CAE3FE91C84E0A94 CRC64; 1;

Query Match 29.2%; Score 52.5; DB 1; Length 508;
Best Local Similarity 34.4%; Pred. NO. 17;
Matches 11; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

Oy 4 SESSTPTGGRFS-GKETPESDDRSORSHME 34
Db ||| | : :||| |: :||| :
442 STSSVEETAAAVSEETPSKDKKKKKKKHGD 473

RESULT 10
ZG20_XENLA STANDARD; PRT; 675 AA.
ID ZG20.XENLA AC PI8714:
AC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95094815; PubMed=8001572;
RA Schaefer U., Rausch O., Boumeester T., Pieler T.;
RT "Sequence-specific recognition of a repetitive DNA element by a C2H2
TL zinc-finger protein in *Xenopus*".;
Eur. J. Biochem. 226:567-576(1994)

[2]
 RN SEQUENCE OF 85-613 FROM N.A.
 RP MEDLINE=90040698; PubMed=2509712;
 RA Macfield W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
 RA Poeling A., Knoechel W.;
 RT "Second-order repeats in Xenopus laevis finger proteins.";
 RL J. Mol. Biol. 208:639-659(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X82843; CAA57965.1; -
 CC PIR: S06565; S06565.
 CC HSSP: P08046; 1A11.
 CC TRANSFAC: T02366; -
 CC InterPro: IPR000822; ZnF_C2H2.
 CC Pfam: PF00096; Zf-C2H2; 18.
 CC PRINTS: PR00048; ZINC_FINGER.
 CC PRODOM: PD000003; ZnF_C2H2; 5.
 CC SMART: SM00355; ZnF_C2H2; 18.
 CC PROSITE: PS00028; ZINC_FINGER_C2H2_1; 17.
 CC PROSITE: PS00157; ZINC_FINGER_C2H2_2; 18.
 CC Zinc-finger; Metal-binding; DNA-binding; Repeat.
 KW 2N_FING 62 84 C2H2-TYPE.
 FT 2N_FING 90 112 C2H2-TYPE.
 FT 2N_FING 118 140 C2H2-TYPE.
 FT 2N_FING 146 168 C2H2-TYPE.
 FT 2N_FING 174 196 C2H2-TYPE.
 FT 2N_FING 202 224 C2H2-TYPE.
 FT 2N_FING 257 279 C2H2-TYPE.
 FT 2N_FING 286 308 C2H2-TYPE.
 FT 2N_FING 344 366 C2H2-TYPE.
 FT 2N_FING 373 395 C2H2-TYPE.
 FT 2N_FING 424 446 C2H2-TYPE.
 FT 2N_FING 452 474 C2H2-TYPE.
 FT 2N_FING 507 529 C2H2-TYPE.
 FT 2N_FING 535 557 C2H2-TYPE.
 FT 2N_FING 563 585 C2H2-TYPE.
 FT 2N_FING 591 613 C2H2-TYPE.
 FT 2N_FING 619 642 C2H2-TYPE.
 SQ SEQUENCE 675 AA; 77116 MW; 033094852C1FCF39 CRC64;
 Query Match 28.9%; Score 52; DB 1; Length 675;
 Best Local Similarity 37.5%; Pred. No. 28;
 Matches 12; Conservative 8; Mismatches 10; Indels 2; Gaps 1;
 Oy 3 QSESSTPGSGKEPSEDROSREHMG 34
 Db 392 QSTHTSPSTE--FGVOTEDNHQSPSKDHGE 421
 RESULT 11
 UBP_HUMAN STANDARD; PRT; 902 AA.
 AC Q9Y4E8; Q9HCA6; (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 15 (EC 3.1.2.15) (Ubiquitin
 DE thiolesterase 15) (ubiquitin-specific processing protease 15)
 DE (Deubiquitinating enzyme 15) (Unp-2).
 GN USP15 OR KIAA0529.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;
 RA Kimura Y., Saya H., Nakao M.;
 RT "Cloning and identification of human Unp-2.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RC MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -I- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O -
 CC ubiquitin + a thiol.
 CC -I- TISSUE SPECIFICITY: Expressed in skeletal muscle, kidney, heart,
 CC placenta, liver, thymus, lung, and ovary, with little or no
 CC expression in other tissues.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF013990; AAG28973.1; -
 CC EMBL: AB011101; BAA25455.1; ALT_INIT.
 CC MEROPS: C19.022; -
 CC GeneW: HGNC:12613; USP15.
 CC MIM: 604731; -
 CC InterPro: IPR001394; UCH-2.
 CC Pfam: PF00442; UCH-1; 1.
 CC Pfam: PF00443; UCH-2; 1.
 CC DR PROSITE: PS00972; UCH_2_1; 1.
 CC DR PROSITE: PS00973; UCH_2_2; 1.
 CC DR PROSITE: PS00335; UCH_2_3; 1.
 KW Ubl conjugation pathway; Hydrolyase; Thiol protease; Multigene family.
 FT ACT_SITE 219 219 BY SIMILARITY.
 FT ACT_SITE 804 804 BY SIMILARITY.
 FT ACT_SITE 812 812 BY SIMILARITY.
 FT CONFLICT 163 166 MISSING (IN REF. 2).
 FT CONFLICT 480 480 A -> T (IN REF. 2).
 FT CONFLICT 889 889 H -> N (IN REF. 2).
 SQ SEQUENCE 902 AA; 103429 MW; 4E526C08B0C16E16 CRC64;
 Query Match 28.9%; Score 52; DB 1; Length 902;
 Best Local Similarity 32.3%; Pred. No. 39;
 Matches 10; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 Oy 3 QSESSTPGSGKEPSEDROSREHMG 33
 Db 576 EMETDEPDDESQDQLPSENENSQSEDSVG 606
 RESULT 12
 KFMS_MOUSE STANDARD; PRT; 977 AA.
 AC P09581; Q9DBH9; (Rel. 10; Created)
 DT 01-MAR-1989 (Rel. 10; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
 DE (EC 2.7.1.112) (Fms proto-oncogene) (c-fms).
 GN CSF1R OR CSFMR OR FMS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.
 RX MEDLINE-88217329; PubMed-2966922;
 RA Rothwell V.M.; Rohrschneider L.R.;
 RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
 RT expression.";
 RL Oncogene Res. 1:311-324(1987).
 RN [2]
 RA REVIEWS.
 RL Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RX MEDLINE-93181280; PubMed-8441691;
 RA de Parseval N., Boredeaux D., Gisselbrecht S., Sola B.;
 RT "Reassessment of the murine c-fms proto-oncogene sequence.";
 RL Nucleic Acids Res. 21:750-750(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX STRAIN-C57BL/6J; TISSUE-Liver;
 RA MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aikawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli J., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gatholdi M.,
 RA Gustingich S., Hill D., Hofmann M., Hume D.A., Kauliy M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE OF 1-16 FROM N.A.
 RX MEDLINE-93268269; PubMed-8497248;
 RA Yue X., Favot P., Dunn T.L., Cassidy A.I., Hume D.A.;
 RT "Expression of mRNA encoding the macrophage colony-stimulating factor
 RT receptor (c-fms) is controlled by a constitutive promoter and tissue-
 RT specific transcription elongation.";
 RL Mol. Cell. Biol. 13:3191-3201(1993).
 RN [6]
 RP AUTOPHOSPHORYLATION SITES.
 RX MEDLINE-90258890; PubMed-2160591;
 RA van der Geer P., Hunter T.;
 RT "Identification of tyrosine 706 in the kinase insert as the major
 RT colony-stimulating factor 1 (CSF-1)-stimulated autophosphorylation
 RT site in the CSF-1 receptor in a murine macrophage cell line.";
 RL Mol. Cell. Biol. 10:2991-3002(1990).
 CC -1- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES
 CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; X06368; CAA29666.1; ALT_SEQ.
 DR EMBL; AK004947; BAB33691.1; -
 DR EMBL; S62219; -; NOT_ANNOTATED_CDS.
 DR PIR; S01880; TWMSMD.
 DR HSP; P11362; JRCG.
 DR MGD; MGI:1339758; Cgflr.
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR001824; RTKinaseIII.
 DR InterPro; IPR001245; Tyr_Pkinase.
 DR Pfam; PF00047; Ig; 4.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Euk_Pkinase; 2.
 DR SMART; SM00410; Ig_Like; 3.
 DR SMART; SM00408; Igc2; 1.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
 KW Immunoglobulin domain; Repeat.
 FT CHAIN 1 977
 FT SIGNAL 1 19
 FT STANAL 20 977
 FT DOMAIN 20 511
 FT TRANSSEM 512 536
 FT DOMAIN 537 977
 FT DOMAIN 24 104
 FT DOMAIN 107 197
 FT DOMAIN 204 298
 FT DOMAIN 299 397
 FT DOMAIN 398 503
 FT DOMAIN 580 913
 FT NP_BIND 586 594
 FT BINDING 614 614
 FT ACT_SITE 776 776
 FT DISULFID 42 84
 FT DISULFID 127 177
 FT DISULFID 224 278
 FT DISULFID 417 483
 FT MOD_RES 697 697
 FT MOD_RES 706 706
 FT MOD_RES 807 807
 FT CARBOHYD 45 45
 FT CARBOHYD 73 73
 FT CARBOHYD 302 302
 FT CARBOHYD 335 335
 FT CARBOHYD 389 389
 FT CARBOHYD 410 410
 FT CARBOHYD 449 449
 FT CARBOHYD 478 478
 FT CARBOHYD 491 491
 FT CONFLICT 446 446
 FT CONFLICT 553 553
 FT CONFLICT 616 616
 FT CONFLICT 744 744
 FT CONFLICT 814 814
 SQ SEQUENCE 977 AA; 109178 MW; 7EDF8310CC98906 CRC64;

Query Match 28.9%; Score 52; DB 1; Length 977;
 Best Local Similarity 31.6%; Pred. No. 43;
 Matches 12; Conservative 7; Mismatches 11; Indels 8; Gaps 1;
 Oy 3 QSESSTSTGSGFSK-----TSEDSDSOSREHM 32
 Db 919 QDYANLPSSGSGSGSGSGSGSGSGSSSEPESSSEHL 956

RESULT 13

KFMS_RAT
ID KFMS_RAT STANDARD; PRT; 978 AA.
AC 000495;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R) (EC 2.7.1.112) (Fms proto-oncogene) (c-fms).
GN CSF1R OR CSFMR OR FMS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Skeletal muscle;
RX MEDLINE=9300125; PubMed=1389227;
RA Borczyk A.G., Guillier M., Leibovitch M.P., Leibovitch S.A.;
RT Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis and regulation during myogenesis.";
RL Growth factors 6:209-218(1992).
CC -1- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN TYROSINE KINASE TRANSMEMBRANE RECEPTOR.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61479; CAA43706.1; -
DR PIR; S16385; S16385.
DR HSP; P11362; 1FGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003506; Ig_MHC.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_1ike.
DR InterPro: IPR001824; RTkinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_4.
DR Pfam; PF00069; Pkinase_1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00410; IG_1ike; 3.
DR SMART; SM00219; TYRKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane; Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 978
FT FT
FT TRANSSEM 20 511
FT DOMAIN 512 536
FT FT
FT DOMAIN 537 978
FT FT
FT DOMAIN 24 104
FT FT
FT DOMAIN 107 197
FT FT
FT DOMAIN 204 298
FT FT
FT DOMAIN 299 397
FT FT
FT DOMAIN 398 503
FT FT
FT DOMAIN 580 914
FT NP_BIND 586 594
FT ATP (BY SIMILARITY).

FT BINDING 614 614
FT ACT_SITE 776 776
FT DISULFID 42 84
FT DISULFID 127 177
FT DISULFID 224 278
FT DISULFID 417 483
FT MOD_RES 697 697
FT MOD_RES 706 706
FT MOD_RES 807 807
FT MOD_RES 807 807
FT CARBOHYD 45 45
FT CARBOHYD 73 73
FT CARBOHYD 302 302
FT CARBOHYD 335 335
FT CARBOHYD 389 389
FT CARBOHYD 410 410
FT CARBOHYD 449 449
FT CARBOHYD 478 478
FT CARBOHYD 491 491
SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;
Query Match 28.98; Score 52; DB 1; Length 978;
Best Local Similarity 31.68; Pred. No. 43;
Matches 12; Conservative 7; Mismatches 11; Indels 8; Gaps 1;
OY 3 QSESPPTGSGFGSGKE-----TPSEDROSREHM 32
DB 919 QDYANLPSSGSGSGSDSGSGSGSSSEPEBSSEHL 956
RESULT 14
CREA ASPNG
ID, CREA ASPNG STANDARD; PRT; 427 AA.
AC Q05620;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-binding protein crea (Carbon catabolite repressor).
GN CREA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9336180; PubMed=8359691;
RA Drysdale M.R., Kolze S.E., Kelly J.M.;
RT "The Aspergillus niger carbon catabolite repressor encoding gene, crea.";
RL Gene 130:241-245(1993).
CC -1- FUNCTION: INVOLVED IN CARBON CATABOLITE REPRESSION. REPRSESSES THE TRANSCRIPTION OF THE ALCA, ALCA AND ALDA GENES BY BINDING TO A GC-RICH REGION IN THEIR PROMOTER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE CREA/MIG GROUP OF C2H2-TYPE ZINC-FINGERS PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03811; AAA32690.1; -
DR PIR; JN0785; JN0785.
DR HSP; P07248; 1ARD.
DR InterPro: IPR000822; znf_C2H2.
DR Pfam; PF00066; zf-C2H2; 2.
DR PRINTS; PR00048; ZINCFINGER.

DR ProDom: PD000003; Znf_C2H2; 1.
 DR SMART: SM00355; Znf_C2H2; 2.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 2.
 KM DNA-binding: Transcription regulation; Repressor; Zinc-finger;
 KM Nuclear protein; Metal-binding
 FT ZN_FING 76 98 C2H2-TYPE.
 FT ZN_FING 104 128 C2H2-TYPE.
 FT DOMAIN 143 150 POLY-ALA.
 FT DOMAIN 275 281 ASP-RICH (ACIDIC).
 SQ SEQUENCE 427 AA; 46141 MW; 4CA3D3B7DD148456 CRC64;

Query Match 28.6%; Score 51.5; DB 1; Length 427;
 Best Local Similarity 46.9%; Pred. No. 19;
 Matches 15; Conservative 2; Mismatches 14; Indels 1; Gaps 1;

QY 4 SESSTPGSGSKETPSDDRSQREHMG 35
 DB 20 STDSTPTPVDSK-TPSTPSSQSNMANS 50

RESULT 15
 NXSL_NAJAT STANDARD; PRT; 83 AA.
 AC P01430; O13079;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Short neurotoxin 1 precursor (Cobrotoxin).
 OS Naja atra (Chinese cobra).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Elapidae; Elapinae; Naja.
 OX NCBI_TaxID=8656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Venom gland;
 RA Chu R.C., Yang C.C.;
 RL Submitted (Mar-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RA MEDLINE=98042466; PubMed=9367842;
 RX Chang L.-S., Lin J., Chou Y.-C., Hong E.;
 RT "Genomic structures of cardiotoxin 4 and cobrotoxin from Naja naja atra (Taiwan cobra).";
 RL Biochem. Biophys. Res. Commun. 239:756-762(1997).
 RN [3]
 RP SEQUENCE OF 22-83.
 RC TISSUE-Venom;
 RA MEDLINE=69294410; PubMed=5820687;
 RX Yang C.C., Yang H.J., Huang J.S.;
 RT "The amino acid sequence of cobrotoxin.";
 RL Biochim. Biophys. Acta 188:65-77(1969).
 RN [4]
 RP SEQUENCE OF 22-83.
 RC TISSUE-Venom;
 RA MEDLINE=93192268; PubMed=8448165;
 RX Chlou S.H., Raynor R.L., Zheng B., Chambers T.C., Kuo J.F.;
 RT "Cobra venom cardiotoxin (cytotoxin) isoforms and neurotoxin: comparative potency of protein kinase C inhibition and cancer cell cytotoxicity and modes of enzyme inhibition.";
 RL Biochemistry 32:2062-2067(1993).
 RN [5]
 RP DISULFIDE BONDS.
 RX MEDLINE=71125117; PubMed=5533659;
 RA Yang C.C., Yang H.J., Chiu R.H.C.;
 RT "The position of disulfide bonds in cobrotoxin.";
 RL Biochim. Biophys. Acta 214:353-363(1970).
 RN [6]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91065386; PubMed=2249693;
 RA Yu C., Lee C.-S., Chuang L.-C., Shel Y.-R., Wang C.Y.;

RT "Two-dimensional NMR studies and secondary structure of cobrotoxin in aqueous solution.";
 RL Eur. J. Biochem. 193:789-799(1990).
 RN [7]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93183875; PubMed=8443154;
 RA Yu C., Bhaskaran R., Chuang L.-C., Yang C.C.;
 RT "Solution conformation of cobrotoxin: a nuclear magnetic resonance and hybrid distance geometry-dynamics simulated annealing study.";
 RL Biochemistry 32:2131-2136(1993).
 CC -1- FUNCTION: PRODUCES PERIPHERAL PARALYSIS BY BLOCKING NEUROMUSCULAR TRANSMISSION AT THE POSTSYNAPTIC SITE. BINDS TO THE NICOTINIC ACETYLCHOLINE RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- MISCELLANEOUS: LD50) IS 0.09 MG/KG BY SUBCUTANEOUS INJECTION.
 CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U58519; AAB03221.1; -
 CC EMBL: U58520; AAB03222.1; -
 CC EMBL: U58521; AAB03223.1; -
 CC EMBL: U77490; AAB36930.1; -
 CC EMBL: U77491; AAB36931.1; -
 CC EMBL: U77492; AAB36932.1; -
 CC EMBL: U42582; AAB01538.1; -
 CC EMBL: Y12492; CAA73097.2; -
 CC PIR: A01698; NINJIF.
 CC PIR: D40667; D40667.
 CC PDB: 1COD; 20-JUL-95.
 CC PDB: 1COE; 20-JUL-95.
 CC InterPro: IPR003571; Snake_toxin.
 CC DR Pfam: PF00087; toxin; 1.
 CC DR ProDom: PD000206; Snake_toxin; 1.
 CC DR PROSITE: PS00272; SNAKE_TOXIN; 1.
 CC KM Venom; Neurotoxin; Postsynaptic neurotoxin; Multigene family; signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 83 SHORT NEUROTOXIN 1.
 FT DISULFID 24 45
 FT DISULFID 38 62
 FT DISULFID 64 75
 FT DISULFID 76 81
 SQ SEQUENCE 83 AA; 9261 MW; 4DD6077C92717052 CRC64;

Query Match 28.3%; Score 51; DB 1; Length 83;
 Best Local Similarity 41.9%; Pred. No. 3.4;
 Matches 13; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 3 QSSSTPGSGSKETPSDDRSQREHMG 33
 DB 27 QSSSTPTTTCGSGGETNCT--KKRWRDHG 55

Search completed: June 17, 2003, 11:56:47
 Job time : 4.10102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:51:11 ; Search time 10.205 Seconds

(without alignments)
827.825 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDDERSTDSQCCSEDEI.....ETTAQVSPPRGKRWRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SPTREMBL_21.*

1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteria:*
17: sp._archaea:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	1807	4	Q9ULH6
2	195.5	90.5	592	4	Q96CH9
3	66	30.6	584	11	Q9CR92
4	65.5	29.4	521	4	Q9Y528
5	63.5	29.2	1114	5	Q9NED2
6	63	29.2	707	11	Q99K50
7	62.5	28.9	642	5	Q9V7M4
8	62.5	28.9	652	5	Q8T092
9	62	28.7	422	4	Q8W0N3
10	61	28.2	1347	10	Q38969
11	61	28.2	1854	10	Q9C606
12	61	28.2	2599	10	Q9C727
13	60.5	28.0	715	11	Q9C2X1
14	60	27.8	484	11	Q9CS98
15	60	27.8	1144	5	Q9N0H7
16	60	27.8	1657	10	Q8W0X4

17	59.5	27.5	292	10	Q9LVH1	Q9LVH1 arabidopsis
18	59.5	27.5	413	4	Q96AJ1	Q96AJ1 homo sapien
19	59.5	27.5	1142	5	Q95XD2	Q95XD2 caenorhabditis
20	59	27.3	586	16	Q9H2D4	Q9H2D4 pseudomonas
21	59	27.3	957	4	Q9BXT9	Q9BXT9 homo sapien
22	59	27.3	1306	11	Q923M3	Q923M3 mus musculu
23	58.5	27.1	182	4	Q9NSU7	Q9NSU7 homo sapien
24	58.5	27.1	517	4	Q96L79	Q96L79 homo sapien
25	58.5	27.1	541	4	Q96J78	Q96J78 homo sapien
26	58.5	27.1	566	4	Q9UP68	Q9UP68 homo sapien
27	58.5	27.1	582	13	P79741	P79741 brachydanio
28	58	26.9	1520	4	Q15087	Q15087 homo sapien
29	58	26.9	1781	4	Q9UKX0	Q9UKX0 homo sapien
30	58	26.9	1784	6	Q8WML3	Q8WML3 macaca fasc
31	58	26.9	1890	4	Q9UKW3	Q9UKW3 homo sapien
32	58	26.9	2072	4	Q8WYB5	Q8WYB5 homo sapien
33	58	26.9	2073	4	Q9UKW2	Q9UKW2 homo sapien
34	58	26.9	4065	3	Q9P421	Q9P421 neurospora
35	57.5	26.6	88	16	Q931C4	Q931C4 rhizobium m
36	57	26.4	339	10	Q8RWE0	Q8RWE0 arabidopsis
37	57	26.4	340	10	Q9LEF1	Q9LEF1 arabidopsis
38	57	26.4	396	5	Q04151	Q04151 toxoplasma
39	57	26.4	513	10	Q9ARE4	Q9ARE4 capsella ru
40	57	26.4	539	5	Q9VPM5	Q9VPM5 drosophila
41	57	26.4	548	11	Q91XD1	Q91XD1 mus musculu
42	57	26.4	603	12	Q81076	Q81076 human para
43	57	26.4	932	5	Q9VZP5	Q9VZP5 drosophila
44	57	26.4	954	5	Q9U0M5	Q9U0M5 plasmodium
45	57	26.4	2643	5	Q01352	Q01352 caenorhabditis

ALIGNMENTS

RESULT 1

ID Q9ULH6 PRELIMINARY: PRT: 1807 AA.
AC Q9ULH6: Q96P46;
DT 01-MAY-2000 (TREMUREL. 13, Created)
DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE KIAA1244 protein (BIG3) (Fragment).
GN KIAA1244.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [2]
RP SEQUENCE OF 38-1807 FROM N.A.
RA Hong W.;
RT "KIAA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
RT subfamily of Arf GEFs.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB033070; BAA6558.1;
DR EMBL: AF413080; AL04174.1;
DR InterPro: IPR000904; Sec7.
DR SMART: SM00222; Sec7; 1.
FT NON_TER
SQ SEQUENCE 1807 AA; 199921 MW; B762C29916F72CB1 CRC64;

Query Match 100.0%; Score 216; DB 4; Length 1807;
Best Local Similarity 100.0%; Pred. NO. 6.8e-20;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWRAR 41
 DB 1473 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWRAR 1513

RESULT 2

O96CH9 PRELIMINARY; PRT; 592 AA.
 AC O96CH9; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 66.4 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RL Strausberg R.;
 DR EMBL: BC014227; A4114227.1; -
 KM Hypothetical protein. 1
 FT NON_TER 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match 90.5%; Score 195.5; DB 4; Length 592;
 Best Local Similarity 97.5%; Pred. No. 1.1e-17;
 Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWRAR 40
 DB 259 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGK-KRQWRA 297

RESULT 3
 O9CR92 PRELIMINARY; PRT; 584 AA.
 AC O9CR92; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 4921513E08Rik protein.
 GN 4921513E08Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Aoechi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
 Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 Suzuki H., Togo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).

DR EMBL: AK015646; BAB29914.1; -
 DR EMBL: AK014878; BAB29599.1; -
 DR MGD: MGI:1913967; 4921513E08Rik.
 SQ SEQUENCE 584 AA; 67255 MW; 0F360DF474D78E18 CRC64;

Query Match 30.6%; Score 66; DB 11; Length 584;
 Best Local Similarity 35.9%; Pred. No. 1.7;
 Matches 14; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

OY 1 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWR 39
 DB 125 DEDEDEDEDEDEDEDEDEDEDEGERDRPKRCKRCKR 163

RESULT 4

O9Y5Z8 PRELIMINARY; PRT; 521 AA.
 AC O9Y5Z8; 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Male-specific lethal-3 homolog 1.
 GN MSL3L1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9936525; Pubmed=10395802;
 RA Prakash S.K., Van den Veyver I.B., Franco B., Volta M., Ballabio A.,
 Zoghbi H.Y.;
 RT Characterization of a novel chromo domain gene in Xp22.3 with
 RT homology to Drosophila msl-3.
 RL Genomics 59:77-84(1999).
 DR EMBL: AF117065; AAD38499.1; -
 DR InterPro: IPR000953; Chromo.
 DR SMART: SM00298; CHROMO_1.
 SQ SEQUENCE 521 AA; 59807 MW; 0C85AE40FDF874CD CRC64;

Query Match 30.3%; Score 65.5; DB 4; Length 521;
 Best Local Similarity 35.4%; Pred. No. 1.8;
 Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

OY 1 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWRAR 41
 DB 123 EDDERSTDSSQCCSSEDEDFEETAAVSPPRGKCKRQWRAR 170

RESULT 5

O9NED2 PRELIMINARY; PRT; 1114 AA.
 AC O9NED2; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 119.2 kDa protein.
 GN P1105.11.
 OS Leishmania major.
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NC NCB1_Taxid=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
 Rajandream M.A., Barrell B.G.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RX MEDLINE=98146435; Pubmed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 Smith D.F.;

```

RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL139794; CAB75569.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 1114 AA; 119250 MW; DB8A995FA5793A17 CRC64;

Query Match
Best Local Similarity 29.4%; Score 63.5; DB 5; Length 1114;
Matches 13; Conservative 11; Mismatches 15; Indels 7; Gaps 1;

OY 3 DERSTDSOOCCSEDEDEFEEFTAGVSPRGK-----EKRQWRAR 41
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 63 EEVDVVSATTCASSPEAVLEQPRHYAPPOQRGRLQRRPKRWMSR 108

RESULT 6
O99K50 ID O99K50 PRELIMINARY; PRT; 707 AA.
AC O99K50.
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nucleolin-.
GN NCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC005460; AAH05460.1; -.
DR HSSP; P09651; 1HA1.
DR MGD; MGI:97286; NGL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; PS00102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_3.
SQ SEQUENCE 707 AA; 76779 MW; 44B42C6E146E5D50 CRC64;

Query Match
Best Local Similarity 29.2%; Score 63; DB 11; Length 707;
Matches 11; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 1 EDDERSTDSSOCCSEDEDIEFEETNOVAPPGRKE 34
   |||| | :: |::| || :::: ||:
DB 193 EDDEDEDEDEEDDEEDSBEVEWETTANGKK 226

RESULT 7
O9V7N4 ID O9V7N4 PRELIMINARY; PRT; 642 AA.
NC O9V7N4.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CG4282 protein.
GN CG4282.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides F.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
```

RA Brandon N.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abdl J.F., Agbayan A., An H.-J., Andrews-Piankoff C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brötter P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jatalai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liu X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Luo X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskenen D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Paul V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skipski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003807; AAP58014.1;
DR FlyBase; FBgn0034114; CG4282.
DR InterPro; IPR000637; AT_hook.
DR InterPro; IPR000822; Zn_f_C2H2.
DR Pfam; PF00096; zt-C2H2; 9.
DR SMART; SM00384; AT_hook; 1.
DR SMART; SM00355; znf_C2H2; 9.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 7.
KW DNA-binding; Metal-binding; Zinc-finger
SO SEQUENCE 642 AA; 74195 MW; F3B91AB9FB9DFDDE CRC64;

Query Match 28.9%; Score 62.5; DB 5; Length 642;
Best Local Similarity 40.5%; Pred. No.5.6;
Matches 15; Conservative 7; Mismatches 12; Indels 3; Gaps 1;

Oy 5 RSTDSSOCCSSEDDIEFTAAQVSPRCKEKQRMAR 41
Db 160 KAEDSSVDKDKDEDGEVETAEAPPK---KRGRPR 193

RESULT 8
ID O8T092 PRELIMINARY; PTR; 652 AA.
AC O8T092;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD2142IP.
GN CG4282.
OS *Drosophila melanogaster* (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayanl A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Paclob J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,
 RU Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069744; AAL39619.1; -
 DR EMBL: AY069744; AAL39619.1; -
 SQ SEQUENCE 652 AA; 75232 MW; C8724857A3015F87 CRC64;

Query Match 28.9%; Score 62.5; DB 5; Length 652;
 Best Local Similarity 40.5%; Pred. No. 5.7;
 Matches 15; Conservative 7; Mismatches 12; Indels 3; Gaps 1;

OY 5 RSTDSSQCCSEDEDFEETAOVSPRGKKRMRAR 41
 DB 170 KAEDGSSVDKDKDEDEVEETAEAPPK--KRGRRR 203

RESULT 9
 O8MUN3 PRELIMINARY; PRT; 422 AA.

AC O8MUN3; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 49.1 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homindae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CERVIX;
 RL Strausberg R.;
 DR Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC019922; AAI19922.1; -
 DR InterPro: IPR000294; Vitk_dep_GLA.
 DR InterPro: IPR000823; ZnF_C2H2.
 DR Pfam: PF00096; zf-C2H2; 8
 DR PRINTS: PRO0046; ZINC_FINGER.
 DR PRODOM: PD000003; ZnF_C2H2; 1.
 DR SMART: SM00355; ZnF_C2H2; 8.
 DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_7.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
 KM Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 422 AA; 49143 MW; 2B146F2E574638FB CRC64;

Query Match 28.7%; Score 62; DB 4; Length 422;
 Best Local Similarity 43.8%; Pred. No. 4.2;
 Matches 14; Conservative 6; Mismatches 10; Indels 2; Gaps 1;

OY 10 SQCCSEDEDFEETAOVSPRGKKRMRAR 39
 DB 22 SQRAKSESDHVRNLFKEEMSKTEGKLENCWR 53

RESULT 10
 O38969 PRELIMINARY; PRT; 1347 AA.

AC O38969; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 153.9 kDa protein.
 GN SABRE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=WS ECOTYPE;
 RA MEDLINE=95172383; PubMed=7867930;
 RA Aschbacher R.A., Hauser M.T., Feldmann K.A., Benfey P.N.;
 RT "The SABRE gene is required for normal cell expansion in
 Arabidopsis";
 RL Genes Dev. 9:330-340(1995).

DR EMBL: U19134; AAC49734.1; JOINED.
 DR EMBL: U19121; AAC49734.1; JOINED.
 DR EMBL: U19122; AAC49734.1; JOINED.
 DR EMBL: U19123; AAC49734.1; JOINED.
 DR EMBL: U19124; AAC49734.1; JOINED.
 DR EMBL: U19125; AAC49734.1; JOINED.
 DR EMBL: U19126; AAC49734.1; JOINED.
 DR EMBL: U19127; AAC49734.1; JOINED.
 DR EMBL: U19128; AAC49734.1; JOINED.
 DR EMBL: U19129; AAC49734.1; JOINED.
 DR EMBL: U19130; AAC49734.1; JOINED.
 DR EMBL: U19131; AAC49734.1; JOINED.
 DR EMBL: U19132; AAC49734.1; JOINED.
 DR EMBL: U19133; AAC49734.1; JOINED.
 KM Hypothetical protein.
 SQ SEQUENCE 1347 AA; 153905 MW; 0011A29D5D8D0DD0 CRC64;

Query Match 28.2%; Score 61; DB 10; Length 1347;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 9 SSOCCSEDEDFEETAOVSP 29
 DB 716 SSIQCPTEDEDEVEEADVP 736

RESULT 11
 O9C606 PRELIMINARY; PRT; 1854 AA.

AC O9C606; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Hypothetical 210.0 kDa protein (Fragment).
 GN T18124.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Felblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz G.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pat G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana";
 RL Nature 408:816-820(2000).
 DR EMBL: AC079131; AAG50770.1; -
 KM Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 1854 AA; 209958 MW; C2A5FB46A994C94C CRC64;

Query Match 28.2% Score 61; DB 10; Length 1854;
 Best Local Similarity 57.1% Pred. No. 28;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 9 SSQCCSEDEDFEETAQVSP 29
 |||:||||:|:|
 Db 1221 SSQCCPEDEDEVEEADEVP 1241

RESULT 12

09C727 PRELIMINARY; PRT; 2599 AA.
 AC 09C727;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE Hypothetical 292.4 kDa protein.
 GN FlgM22.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 CC NCBI_TaxID=3702;
 RX MEDLINE=21016719; PubMed=11130712;
 RP STRAIN=CV. COLUMBIA.
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Miltcher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL: AC073943; AAC50951.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 2599 AA; 292445 MW; 6FAF3CCA592FAA3C CRC64;

Query Match 28.2% Score 61; DB 10; Length 2599;
 Best Local Similarity 57.1% Pred. No. 40;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 9 SSQCCSEDEDFEETAQVSP 29
 |||:||||:|:|
 Db 1968 SSQCCPEDEDEVEEADEVP 1988

RESULT 13

09Q2X1 PRELIMINARY; PRT; 715 AA.
 AC 09Q2X1;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Nucleolin-related protein NRP.
 GN NRP.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=99419025; PubMed=10488083;
 RA Sorokina E.A., Kleiman J.G.;
 RT "Cloning and preliminary characterization of a calcium-binding protein
 closely related to nucleolin on the apical surface of inner medullary
 collecting duct cells.";
 RL J. Biol. Chem. 274:27491-27496(1999).
 DR EMBL: AF151373; AA05625.1; -;
 DR HSSP: P09651; 1HA1.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 4.
 DR SMART: SM00360; RRM; 4.
 DR PROSITE: PS50102; RRM; 4.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN.3.
 SQ SEQUENCE 715 AA; 77439 MW; 54A20FADBAC6987 CRC64;

Query Match 28.0% Score 60.5; DB 11; Length 715;
 Best Local Similarity 37.8% Pred. No. 12;
 Matches 14; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

Qy 1 EDDERSTDSQCCSEDE---DFEETAQVSPRGKE 34
 ||||| | ||| | | : | : |||
 Db 186 EDDEDEDEDEDEDEDEDEDESEEAETTPARGKK 222

RESULT 14

09CS98 PRELIMINARY; PRT; 484 AA.
 AC 09CS98;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 3930402D05RIK protein (fragment).
 GN 3930402D05RIK.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata J., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK014457; BAB29364.1; -;
 DR MGD: MGI:1921269; 3930402D05RIK.
 FT NON_TER 1 1

Qy SEQUENCE 484 AA; 55413 MW; 328371FF52141F54 CRC64;
 Best Local Similarity 29.3% Pred. No. 8.9;
 Matches 12; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

Query Match 27.8% Score 60; DB 11; Length 484;
 Best Local Similarity 29.3% Pred. No. 8.9;
 Matches 12; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 11:51:11 ; Search time 9.95608 Seconds
(without alignments)
827.825 Million cell updates/sec

Title: US-09-991-681-30
Perfect score: 215
Sequence: 1 SPKVEKKDPSRRKKEMWENAGNKIYTMAADKTISKLMTEYK 40

Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	592	4	Q96CH9
2	215	100.0	1807	4	Q9ULH6
3	65	30.2	177	5	Q9BUE8
4	64	29.8	621	2	Q59301
5	63.5	29.5	604	3	Q74491
6	61	28.4	3306	10	Q9FT44
7	59	27.4	277	16	Q8U805
8	38	27.0	672	10	Q941M7
9	57.5	26.7	921	16	Q92926
10	56.5	26.3	201	16	Q8XB83
11	56.5	26.3	483	10	Q64951
12	56.5	26.3	574	10	Q64950
13	56	26.0	304	10	Q9S1S6
14	56	26.0	328	10	Q8S341
15	56	26.0	408	2	Q9L8P3
16	55.5	25.8	103	17	Q8TV49

17	55.5	25.8	676	5	Q9N4T5	Q9N4T5 caenorhabd
18	55	25.6	177	10	Q24127	Q24127 nicotiana t
19	55	25.6	351	10	Q9SMC1	Q9SMC1 nicotiana r
20	55	25.6	351	10	Q9SMB7	Q9SMB7 nicotiana t
21	55	25.6	351	10	Q9MB29	Q9MB29 nicotiana t
22	55	25.6	634	10	Q9FLI2	Q9FLI2 arabidopsi
23	54.5	25.3	555	10	Q9AV96	Q9AV96 hordeum vul
24	54.5	25.3	673	10	Q941M6	Q941M6 pseudomonas
25	54	25.1	732	2	Q8VP74	Q8VP74 oryza sativ
26	54	25.1	1048	10	Q04938	Q04938 rhodococcus
27	53.5	24.9	331	2	Q93EX6	Q93EX6 oryza sativ
28	53.5	24.9	1055	5	Q9LTV9	Q9LTV9 oryza sativ
29	53	24.7	176	5	Q9BUE0	Q9BUE0 plasmidum
30	53	24.7	186	4	Q9P0X1	Q9P0X1 homo sapien
31	53	24.7	259	5	Q93905	Q93905 caenorhabd
32	53	24.7	373	10	Q8RUE5	Q8RUE5 oryza sativ
33	53	24.7	410	10	Q9S2X1	Q9S2X1 arabidopsi
34	53	24.7	423	10	Q9EK39	Q9EK39 arabidopsi
35	53	24.7	462	16	Q07732	Q07732 mycobacteri
36	53	24.7	462	16	Q8VJU6	Q8VJU6 mycobacteri
37	53	24.7	488	5	Q9VFE6	Q9VFE6 drosophila
38	53	24.7	528	4	Q9NUQ3	Q9NUQ3 homo sapien
39	53	24.7	729	5	Q21136	Q21136 caenorhabd
40	53	24.7	900	10	Q9FK71	Q9FK71 arabidopsi
41	52.5	24.4	234	10	Q9S7C4	Q9S7C4 hevea brasl
42	52.5	24.4	769	3	Q8XIE4	Q8XIE4 neurospora
43	52.5	24.4	895	6	Q9BDV9	Q9BDV9 lepus capen
44	52.5	24.4	1289	16	Q9W287	Q9W287 thermocoga
45	52	24.2	196	11	Q9D068	Q9D068 mus musculu

ALIGNMENTS

RESULT 1
ID Q96CH9 PRELIMINARY: PRT: 592 AA.
AC Q96CH9: TISSUE-COLON;
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Hypothetical 66.4 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014227; AAH14227.1;
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 592 AA; 66400 MW; FA1EB07B0DE47B5 CRC64;

Query Match 100.0%; Score 215; DB 4; Length 592;
Best local similarity 100.0%; Pred. NO. 6e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPKVEKKDPSRRKKEMWENAGNKIYTMAADKTISKLMTEYK 40
DB 406 SPKVEKKDPSRRKKEMWENAGNKIYTMAADKTISKLMTEYK 445

RESULT 2

ID Q9ULH6 PRELIMINARY: PRT: 1807 AA.
AC Q9ULH6: 096P46;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE KIAA1244 protein (BIG3) (Fragment).

GN KIAA1244.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirotsawa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [2]
RP SEQUENCE OF 38-1807 FROM N.A.
RA Hong W.;
RT KIAA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
RT subfamily of Arp GEFs.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB033070; BAA86558.1; -;
DR EMBL; AF413080; AL04174.1; -;
DR InterPro: IPR000904; Sec7.
DR SMART: SM00222; Sec7; 1.
FT NON_TER 1
SQ SEQUENCE 1807 AA; 19921 MW; B762C3916F72CB1 CRC64;
Query Match 100.0%; Score 215; DB 4; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2.1e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPVEKKDPSRKKEWENAGNKIYTMAADTKISKLTMEYK 40
Db 1621 SPVEKKDPSRKKEWENAGNKIYTMAADTKISKLTMEYK 1660
RESULT 3
OY 09BUE8 PRELIMINARY; PRT; 177 AA.
ID 09BUE8
AC 09BUE8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=701;
RX MEDLINE=21136462; PubMed=11237850;
RA Knäutleb A., Kun J., Deloron P., Krensner P.G., Klinkert M.O.;
RT "Variants of Plasmodium falciparum Erythrocyte Membrane Protein 1
RT Expressed by Different Placental Parasites are Closely Related and
RT Adhere to Chondroitin Sulfate A.";
RL J. Infect. Dis. 183:1165-1169(2001).
DR EMBL; AF334805; AAK28128.1; -;
FT NON_TER 1
SQ SEQUENCE 177 AA; 21207 MW; 318AC7FEFF0102A2 CRC64;
Query Match 30.2%; Score 65; DB 5; Length 177;
Best Local Similarity 52.6%; Pred. No. 0.89;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
OY 6 KDPSPRKKEWENAGNKIY 24
Db 112 KDDTKRKKEWENAGNKIY 130
RESULT 4
OY 059301 PRELIMINARY; PRT; 621 AA.
ID 059301

AC Q59301;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Endo-beta-1,4-xylanase precursor (EC 3.2.1.8).
OS Cellvibrio mixtus.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Cellvibrio.
OX NCBI_TaxID=39650;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIXTUS;
RX MEDLINE=96077124; PubMed=7492333;
RA Millward-Sadler S.J., Davidson K., Hazlewood G.P., Black G.W.,
RA Gilbert H.J., Clarke J.H.;
RT "Novel cellulose-binding domains, NodB homologues and conserved
RT modular architecture in xylanases from the aerobic soil bacteria
RT Pseudomonas fluorescens subsp. cellulosa and Cellvibrio mixtus.";
RL Biochem. J. 312:39-48(1995).
DR EMBL; Z48926; CAA88762.1; -;
DR HSSP; P14768; ICLX.
DR InterPro: IPR005088; CBM_15.
DR InterPro: IPR001000; Glyco_hydro_10.
DR Pfam; PF03426; CBM_15; 1.
DR Pfam; PF00331; Glyco_hydro_10; 1.
DR PRINTS; PR00134; GLYDRLASE10.
DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; UNKNOWN_1.
KM Glycosidase; Hydrolase; Signal; Xylan degradation.
FT SIGNAL 1
FT POTENTIAL 19
SQ SEQUENCE 621 AA; 64929 MW; 830E5B959DC3AB8 CRC64;
Query Match 29.8%; Score 64; DB 2; Length 621;
Best Local Similarity 43.2%; Pred. No. 4.9;
Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;
OY 2 PKVEKKDPSRKKEWENAGNKIYTMAADTKISKLTME 38
Db 175 PIVQIKGSTPGEMGCMAGNELFTAGEDATISCTVTE 211
RESULT 5
OY 074491 PRELIMINARY; PRT; 604 AA.
ID 074491
AC 074491; P78908;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative ATP-dependent RNA helicase C285.03.
GN SPC285.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Seeger K., Harris D., Lyne M., Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 115-604 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RX Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
RN [3]
CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.
DR EMBL; AL031545; CAA20842.1; -;
DR EMBL; D89259; BAA13920.1; -;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD box.
DR InterPro: IPR001650; Helicase_C.


```

DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Hypothetical protein; ATP-binding; Helicase; RNA-binding.
FT NP_BIND 180 187 ATP (POTENTIAL).
FT SITE 292 295 DEAD BOX.
FT CONFLICT 120 120 V -> F (IN REF. 2).
FT CONFLICT 123 123 N -> Y (IN REF. 2).
FT CONFLICT 401 401 F -> L (IN REF. 2).
FT CONFLICT 482 482 R -> W (IN REF. 2).
FT CONFLICT 502 504 DVA -> HFS (IN REF. 2).
FT CONFLICT 516 516 L -> F (IN REF. 2).
FT CONFLICT 527 527 R -> P (IN REF. 2).
FT CONFLICT 557 558 RR -> KK (IN REF. 2).
FT CONFLICT 578 579 VA -> FP (IN REF. 2).
FT CONFLICT 591 591 Y -> C (IN REF. 2).
FT CONFLICT 596 596 Y -> C (IN REF. 2).
SQ SEQUENCE 604 AA; 68344 MW; 099005F2FD2CCD03 CRC64;

Query Match 29.5%; Score 63.5; DB 3; Length 604;
Best Local Similarity 41.2%; Pred. No. 5.6;
Matches 14; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

Oy 2 PKVEKKDPSRRKKEMWENAGNKIYTMADKTSIKL 35
Db 13 PKIEKKS-KRKKRWLMDENKTHVTASEAIERL 45

RESULT 6
O9FT44 PRELIMINARY; PRT; 3306 AA.
AC 09FT44:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE VPS13-1-like protein.
GN ATSG24740.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bryan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL392145; CAC08248.1; -
DR InterPro; IPR001130; TatD_DNase.
DR PROSITE; PS01090; TATD_2; UNKNOWN.1.
SQ SEQUENCE 3306 AA; 368483 MW; 122B0D048E6C1E99 CRC64;

Query Match 28.4%; Score 61; DB 10; Length 3306;
Best Local Similarity 34.4%; Pred. No. 81;
Matches 11; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 SPKVEKKDPSRRKKEMWENAGNKIYTMADKTI 32
Db 203 STLSRKPCKMOKIMHYAONSVLSDSIDKSI 234

RESULT 7
O8U805 PRELIMINARY; PRT; 277 AA.
AC 08U805:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Oxidoreductase.
GN YAFB OR AFU4294 OR AGR.L.1141.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_Taxid=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D., St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavir T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphumachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gattling S., Miller N., Blanchard M.,
RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009357; AAU45088.1; -
DR EMBL; AE008255; AAR69145.1; -
KW Complete proteome.
SQ SEQUENCE 277 AA; 29836 MW; 26E1BBE1C0914AA4 CRC64;

Query Match 27.4%; Score 59; DB 16; Length 277;
Best Local Similarity 37.0%; Pred. No. 9.1;
Matches 10; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Oy 13 KEMWENAGNKIYTMADKTSIKLMEY 39
Db 72 KWMWYGHDAFKSVDSLKRLKTYD 98

RESULT 8
O94IM7 PRELIMINARY; PRT; 672 AA.
AC 094IM7:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE P-type ATPase (Fragment).
GN CaC.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_Taxid=4513;
RN [1]
RP SEQUENCE FROM N.A.
RA Garcia-Deblas B.;
RT "Plant cells express several stress calcium ATPases but apparently no sodium ATPase."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ310845; CAC40033.1; -
DR InterPro; IPR000661; H/K_Nb/K_ATPase.
DR InterPro; IPR001454; Hlgase/hydrilase.
DR Pfam; PF00689; Cation_ATPase_C; 1.

```

Dt	01-JUN-2002 (TREMBLrel_21, last annotation update)
De	Purative serine acetyltransferase of prophage CP-933T.
Cn	Z2984 OR ECS2638
Os	Escherichia coli O157:H7.
Oc	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Ox	Escherichia.
Rn	NCBI_TaxID=83334;
Rp	[1]
Rc	SEQUENCE FROM N.A.
Rl	STRAIN-O157:H7 / EDL933 / ATCC 700927;
Rm	MEDLINE=21074935; PubMed=11206551;
Ra	Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,
Ra	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Ra	Posifal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Ra	Grobeck E.J., Davis N.W., Lam A., Dimantanta E.T., Potamoukis K.,
Ra	Apodaca T., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Ra	Weich R.A., Blattner F.R.;
Rt	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
Rl	Nature 409:529-533(2001).
Rn	[2]
Rp	SEQUENCE FROM N.A.
Rc	STRAIN-O157:H7 / RIMD 0509952;
Rl	MEDLINE=21156231; PubMed=11258796;
Ra	Hayashi T., Makino K., Ohnishi K., Kurokawa K., Ishii K., Yokoyama K.,
Ra	Iida C.-G., Ohtsubo E., Nakayama K., Mureta T., Tanaka M., Toke T.,
Ra	Iida T., Takami H., Honda T., Sasakiwa C., Ogasaara N., Yasunaga T.,
Ra	Kuhara S., Shiba T., Hattori M., Shinagawa H.;
Rt	"Complete genome sequence of enterohaemorrhagic Escherichia coli
Rt	O157:H7 and genomic comparison with a laboratory strain K-12".;
Rl	DNA Res. 8:11-22(2001).
Rm	EMBL; AE005413; AAC56916.1; "
Rm	EMBL; AP002559; BAB36061.1; ALT_INIT.
Dr	InterPro: IPR001451; Hexapep.transf.
Dr	Pfam: PF00132; hexapep: 3.
Sw	Transferase: Complete proteome.
So	SEQUENCE 201 AA; 22838 MW; 06E1F5A1972A6CE0 CRC64;
<hr/>	
Qy	Query Match 26.3%; Score 56.5; DB 16; Length 201;
	Best Local Similarity 35.7%; Pred. No. 14;
	Matches 15; Conservative 7; Mismatches 13; Indels 7; Gaps 2;
Db	52 EKKDPSSK-KEMWEN-----GKIITMADKTIISKMTXY 39
	: : :
	52 EKQRNRNFFFWRLANEMYINGKLHKRAAKKINSKIINF 93
<hr/>	
Result	11
ID	064951 PRELIMINARY; PRT; 483 AA.
AC	064951;
Dt	01-AUG-1998 (TREMBLrel_07, Created)
Dt	01-AUG-1998 (TREMBLrel_07, Last sequence update)
Dt	01-DEC-2001 (TREMBLrel_19, Last annotation update)
De	3-hydroxy-3-methylglutaryl coenzyme A reductase.
Os	Tagetes erecta (African marigold).
Oc	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Oc	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
Oc	Asteridae; Gnasteris II; Asterales; Asteraceae; Asteroideae;
Oc	Helenieae; Tagetes.
Ox	NCBI_TaxID=13708;
Rn	[1]
Rp	SEQUENCE FROM N.A.
Ra	del Villar Martinez A.A., Ma R., Paredes-Lopez O.;
Rt	"Molecular Characterization of cDNAs Encoding 3-Hydroxy-3-
Rt	methylglutaryl Coenzyme A Reductase in Marigold (Tagetes erecta).";
Rl	J. Plant Physiol. 155:205-211(1999).
DR	EMBL; AF034761; AAC15476.1; "
DR	InterPro: IPR002202; HMGC-CoA_red.
DR	Pfam: PF00368; HMGC-CoA_red: 1.
DR	PROSITE: PSS0066; HMGC_COA_REDUCTASE_1; 1.
DR	PROSITE: PSS0065; HMGC_COA_REDUCTASE_4; 1.
SO	SEQUENCE 483 AA; 52816 MW; 850DAEA30AD9AE70 CRC64;

Ox	NCHI_TaxID=727;
Rn	[1]
Rp	SEQUENCE FROM N.A.
Rc	STRAIN=EAGAN;
Rx	MEDLINE-20231799; PubMed-10768954; Chang C.-C., Gilsdorf J.R., Dirita V.J., Mats C.F.; RA "Identification and genetic characterization of Haemophilus influenzae RT genetic island 1."; Rl infect. Immun. 68:2630-2637(2000). DR EMBL: AF198256; AAF27347.1; -;-:- DR InterPro: IPR002104; Phage_integrase. DR Pfam: PF00589; Phage_integrase_1 SO SEQUENCE 408 AA; 46884 MW; 04733BD96C4877F0 CRC64;
Qy	Query Match 26.0%; Score 56; DB 2; Length 408; Best Local Similarity 30.3%; Pred.No. 35; Matches 10; Conservative 9; Mismatches 14; Indels 0; Gaps 0;
Dd	120 KEKAKNPETREKNWERLKNIHLEPYIGDKHVSEI 152 <div style="font-family: monospace;"> : ! : : : : : : : : 3 KVEKKDPSRRKEWENGNKIYTMADDTISKL 35</div>

Search completed: June 17, 2003, 11:59:48
Job time : 11.9561 secs

```

best local similarity 100.00; r100: no: 0.70 220;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RIRAAAOVFMIDTCCSPKTPNNPDHAAOQCIIILPDEKPNHGTKKSVSFREIVSL 60
DB 1230 RIRAAAOVFMIDTCCSPKTPNNPDHAAOQCIIILPDEKPNHGTKKSVSFREIVSL 1349
QY 61 SHOVLQWLYDILLEEFGKSPGPEKTIQVPEAKLAGLRISQNLAVIFDILLDSYR 120
DB 1330 SHOVLQWLYDILLEEFGKSPGPEKTIQVPEAKLAGLRISQNLAVIFDILLDSYR 1409
QY 121 TAREFDTSPGLKCLLKVKVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQVK 180
DB 1410 TAREFDTSPGLKCLLKVKVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQVK 1469
QY 181 VLFEDESDSTSSQCCSEDEDIFFETAOVSPPRCKEKORARARPLLVSQVSNADWV 240
DB 1470 VLFEDESDSTSSQCCSEDEDIFFETAOVSPPRCKEKORARARPLLVSQVSNADWV 1529
QY 241 LVKRLHKLKLMELCNMYIOMHLDLENCMEEPPIFKDPPFILPFSQESSTPTSGFSKGE 300
DB 1530 LVKRLHKLKLMELCNMYIOMHLDLENCMEEPPIFKDPPFILPFSQESSTPTSGFSKGE 1589
QY 301 TPSEDRSQSRHMGESLSLAKAGGDLPLPSPKVEKKDPSRKKEMENAGNKITYTMAD 360
DB 1590 TPSEDRSQSRHMGESLSLAKAGGDLPLPSPKVEKKDPSRKKEMENAGNKITYTMAD 1649
QY 361 KTSIKLMTYKKRKQOHNLSAFPKVEKKGEPLGPGGOSPLLQROHLMDOGOMRHS 420
DB 1650 KTSIKLMTYKKRKQOHNLSAFPKVEKKGEPLGPGGOSPLLQROHLMDOGOMRHS 1709
QY 421 FSAGPELLRODKRPRSGSTGSSLSVSDAEAOIOAMTNMYLVLTNOIQLIPDOTFTALQ 480
DB 1710 FSAGPELLRODKRPRSGSTGSSLSVSDAEAOIOAMTNMYLVLTNOIQLIPDOTFTALQ 1769
QY 481 PAVFPCISQLTCHVTDIRVOAVREMLGRVGRVYDIIV 518
DB 1770 PAVFPCISQLTCHVTDIRVOAVREMLGRVGRVYDIIV 1807

```

RESULT 2

```

Q96CH9 PRELIMINARY; PRT: 592 AA.
AC 096CH9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 66.4 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014227; AAH14227.1;
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 592 AA: 66400 MM; FA1E807B0DF47B5 CRC64;

```

Query Match 99.2%; Score 2675.5; DB 4; Length 592;
 Best Local Similarity 99.6%; Pred. No. 5.5e-217;
 Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 RIRAAAOVFMIDTCCSPKTPNNPDHAAOQCIIILPDEKPNHGTKKSVSFREIVSL 60
DB 1230 RIRAAAOVFMIDTCCSPKTPNNPDHAAOQCIIILPDEKPNHGTKKSVSFREIVSL 1349
QY 61 SHOVLQWLYDILLEEFGKSPGPEKTIQVPEAKLAGLRISQNLAVIFDILLDSYR 120
DB 1330 SHOVLQWLYDILLEEFGKSPGPEKTIQVPEAKLAGLRISQNLAVIFDILLDSYR 1409
QY 121 TAREFDTSPGLKCLLKVKVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQVK 180
DB 1410 TAREFDTSPGLKCLLKVKVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQVK 1469

```

```

DB 196 TAREFDTSPGLKCLLKVKVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQVK 255
QY 181 VLFEDESDSTSSQCCSEDEDIFFETAOVSPPRCKEKORARARPLLVSQVSNADWV 240
DB 256 VLFEDESDSTSSQCCSEDEDIFFETAOVSPPRCKEKORARARPLLVSQVSNADWV 314
QY 241 LVKRLHKLKLMELCNMYIOMHLDLENCMEEPPIFKDPPFILPFSQESSTPTSGFSKGE 300
DB 315 LVKRLHKLKLMELCNMYIOMHLDLENCMEEPPIFKDPPFILPFSQESSTPTSGFSKGE 374
QY 301 TPSEDRSQSRHMGESLSLAKAGGDLPLPSPKVEKKDPSRKKEMENAGNKITYTMAD 360
DB 375 TPSEDRSQSRHMGESLSLAKAGGDLPLPSPKVEKKDPSRKKEMENAGNKITYTMAD 434
QY 361 KTSIKLMTYKKRKQOHNLSAFPKVEKKGEPLGPGGOSPLLQROHLMDOGOMRHS 420
DB 435 KTSIKLMTYKKRKQOHNLSAFPKVEKKGEPLGPGGOSPLLQROHLMDOGOMRHS 494
QY 421 FSAGPELLRODKRPRSGSTGSSLSVSDAEAOIOAMTNMYLVLTNOIQLIPDOTFTALQ 480
DB 495 FSAGPELLRODKRPRSGSTGSSLSVSDAEAOIOAMTNMYLVLTNOIQLIPDOTFTALQ 554
QY 481 PAVFPCISQLTCHVTDIRVOAVREMLGRVGRVYDIIV 518
DB 555 PAVFPCISQLTCHVTDIRVOAVREMLGRVGRVYDIIV 592

```

RESULT 3

```

Q9W444 PRELIMINARY; PRT: 2045 AA.
AC 09W444;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CG5937 protein.
GN CG5937.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
RA Baller R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bounk J., Brokslein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003436; AAF6115.1;
 DR FlyBase: FBgn0029834; CG5937.
 DR InterPro: IPR002173; PFB.
 DR InterPro: IPR000904; Sec7.
 DR SMART: SM00222; Sec7.1.
 DR PROSITE: PS00583; PFB_KINASES_1; UNKNOWN_1
 SO SEQUENCE 2045 AA; 221370 MW; E59DAB2DC13849D7 CRC64;
 Query Match 6.6%; Score 177.5; DB 5; Length 2045;
 Best Local Similarity 29.9%; Pred. No. 1.4e-05;
 Matches 60; Conservative 25; Mismatches 53; Indels 63; Gaps 7;
 QY 1 RIRAAQGVFLDQGP-----KTPNFDHAGSCQILPEDEK-----NHT 46
 DB 1639 RYALAQGVFLSDNQREPGONQAPPS-----ASOCKL-----SDRSTSYFLLYPLNNGFN 1689
 QY 47 KK-----SVSFREIVSLSHOVLLQNLVDLLEEF-----YKGPS 82
 DB 1690 SNLDFVIRIPKKNLYVGLLANOMLQVAKLLSLKLCVQAVSTCFIDNVAASAPSS 1749
 QY 83 PGEKTIQVPEAKLAGFLRYISMONLAVIFDLDDSYFANEDTSPGLKCLKKVSG 142
 DB 1750 HDYDDEFRSKELLFLCVKQYL-----MSALEFDSPLPELKLMOQVSNIE 1793
 QY 143 GAANLYROSASFNIYFHALV 163
 DB 1794 YAAALYKQMTSSMMITYALV 1814
 RESULT 4
 Q9VXT5 PRELIMINARY: PRT; 665 AA.
 AC 09VXT5:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG9213 protein.
 GN CG9213.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan D.A., Bouck J., Brokstein P., Brotlier P.,
 RA Butlis K.C., Busam M.R., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Honck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibgwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclet J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003436; AAF6115.1;
 DR FlyBase: FBgn0030655; CG9213.
 DR InterPro: IPR000904; Sec7.
 DR PROSITE: PS00583; PFB_KINASES_1; UNKNOWN_1
 SO SEQUENCE 665 AA; 75646 MW; 9564807D1BC605F4 CRC64;
 Query Match 4.9%; Score 133; DB 5; Length 665;
 Best Local Similarity 22.1%; Pred. No. 0.016;
 Matches 66; Conservative 43; Mismatches 134; Indels 56; Gaps 12;
 QY 181 VLFEDDERSTDSQCCSEDEDFEETAGVSPRCKEKRRARRAMPILSYOVPSNADWV 240
 DB 4 IOFESGREKDKARQELREAREAMLQQAKEARELROQRORKELR-----GRADWML 54
 QY 241 --LVKRLKLCMELCNNTYQHLDLENMKEERPIFKGDPFILPFOSESTSPGSG 298
 DB 55 PALAKLLEKPAKSKRN--VSRKSKRSKSKSKSKSRKRN-----SSSSSESTSSSF 107
 QY 299 KETPSDDRSOSREHMGESLSLAKAGCD--LLPP--SPVKEKKDPSRKKEWENAGN 352
 DB 108 ----SEDEKERRRRKKKSKRSKESASDENVEAPPLAADNVTYKKEPPQDDW----- 157
 QY 353 KIYMAADTKISKIMTEYKKRRKQOHNLSAFPEKV--EKKGEPLG--RGDSPLL 405
 DB 158 ----MTSLLKTRTSRRKREKPAKPEKA--QIDAYAPKASGRGLNYWKSNGGLPGF 211
 QY 406 QRPQHLMDQGMRSFSAGPELLRODKP-----RSGSTGSSLSVYRDAPAQIO 455
 DB 212 QKPDDDERQAKPHSSSSAOGSSRCGRKPKAKAPSPRRSRKRSKTSATSDDEEEBAVO 270
 RESULT 5
 Q95R14 PRELIMINARY: PRT; 687 AA.
 AC 095R14:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE LD28117P.
 GN CG9213.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX Stapleton M., Brokstein P., Hong L., Agbayani A., Carlsson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Flise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Paclet J., Paragas V., Park S., Pionanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY061353; AAL28901.1;
 DR FlyBase: FBgn0030655; CG9213.
 SO SEQUENCE 687 AA; 78148 MW; 7376BABCE1BE676 CRC64;

Query Match 4.9%; Score 133; DB 5; Length 687;
 Best Local Similarity 22.1%; Pred. No. 0.016;

Matches 66; Conservative 43; Mismatches 134; Indels 56; Gaps 12;

OY 181 VLFEDEESTDSOCCSEDEDIEETAYQVSPRGKREKQRMARAPLLSVQVSNADWVW 240
 DB 4 IQFESGRKKARQELRAREAMLOAKERAEELGQRRQKELR-----GEADWML 54
 OY 241 --LVRLHKLCEMLCNNTYQMHLDLNCMEEPPIFGDPFLLPSFOSESSTPTGFGSG 298
 DB 55 PALKKKLEKPAKKSKKN-VSKHKSRKSKSKSKSKKHNN-----SSSSSESTSSSSSF-- 107
 OY 299 KETPSEDDRSQSRHEHMGESLSLKAGGD---LLLP---SPKVEKKDSRKKEMWENGN 352
 DB 108 ---SEDEKERRKKRKKSKSRKESASDEWEAPPLAUNVTKEPPORDW----- 157
 OY 353 KIYTMADKTIKLTETKRRKQOHNLSAPFEVKV---EKKGEPLG---RGDPSPL 405
 DB 158 ---MTSESLLTKEFSRREKREKAPNEKA---QIDAYPAKSGRELNTYKSNGLPGF 211
 OY 406 QRPQHLMDQGMHRSFSAQPELLRQDKRP-----RSGSTGSSLSVSVDAEAQIQ 455
 DB 212 QKPDDEERQAKPHSSSSAQSGSGRWKPKAKAPSPRRSRKSRKATSSDEEEEAQV 270

RESULT 6

O93R63 PRELIMINARY; PRT: 751 AA.

AC O93R63;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE SD04745P.
 GN G12750.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;

RA SEQUENCE FROM N.A.
 RA Stopleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Xu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY061596; AAL29144.1;
 DR FlyBase: FBgn0032678; CG12750.
 DR InterPro: IPR003891; IF_eIF4G_MA3.
 DR InterPro: IPR003580; Protachykinin.
 DR Pfam: PF02847; MA3; 1.
 DR SMART: SM00203; TK; 2.
 SO SEQUENCE 751 AA; 86453 MW; DEA5672CB2EE71E7 CRC64;

Query Match 4.8%; Score 130.5; DB 5; Length 751;
 Best Local Similarity 22.1%; Pred. No. 0.03;

Matches 89; Conservative 57; Mismatches 157; Indels 99; Gaps 19;

OY 85 EKTIOVEALAGFL--RYISOMLAV--IFDLIDSVRTAREPDTSPGLCKLTKVSG 140
 DB 164 EQRTE-----KFGLLAORPCNINKIYIPPEEIFKQIYOTHRDITN-----RLRNSK 214
 OY 141 IGAANLVROSAMSFNIYFHALVCAVLTNOETTAQV--KKVLFD--DERSTDSSQGCSS 198
 DB 215 F--FAHLFTDAISWDV-----LECIQLNEDDTSSSRIFIKILFQELAEYMGICLANKL 268
 OY 199 EDEDIETFAVSP--PRGKEKRWARMPLLSVQVPSNADWVWLVRLHKLCEMLCN 255

DB 269 KDVVLVESIAGLFPRDNRNTR-----FSINFTSIGLGLTDDLR----- 311
 OY 256 YIQHMLDLENCMEEPPIK-----GDP-----FILPQSESSTPTGFGS 297
 DB 312 -----LKNAPKSVPAINAELANAGNPRDGSAPRGNTKVPASSSSSS-----S 357
 OY 298 GKTPSEDDRSQSRHEHMGESLSLKAGGDLPLPSPKVEKKDPKPK-----KEMWENAG 351
 DB 358 SSDTSDS-----SEEDSSSDSSSESSSDSSSEPKKKRKKRDKKKKATKEKSKTK 414
 OY 352 NKITTMADKTIKLTETKRRKQOHNLSAPFEVKVKEKPEPLGPRQDPELQRPQL 411
 DB 415 NKKKKKRAEK-----EQEKEKQKRSKKEKEDKKRKKKKAANKSKRRKSOES 467
 OY 412 MDQGMHRSFSAQPELLRQDKRPSPSGSTGSSLSVSVDAEAQ 453
 DB 468 SD-----SSGSE--DSDKSTSSSDSSSSSDSEDAERQ 499

RESULT 7

O923M3 PRELIMINARY; PRT: 1306 AA.

AC O923M3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE BM14504.3 (Novel protein similar to human KIAA1210) (Fragment).
 GN BM14504.3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RA SEQUENCE FROM N.A.
 RA Blakey S.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL450395; CAC42194.1;
 DR InterPro: IPR002106; AALRNA_1lgaseII.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 FT NON-TER 1 1306
 FT NON-TER 1306 1306
 SO SEQUENCE 1306 AA; 144102 MW; 6255CFD233456C19 CRC64;

Query Match 4.8%; Score 128.5; DB 11; Length 1306;
 Best Local Similarity 19.3%; Pred. No. 0.098;
 Matches 90; Conservative 51; Mismatches 123; Indels 203; Gaps 20;

OY 184 EDDERSTDSQGC-----SDEDEDIEETAYQVSPRGKREKQRMARAPLLSVQVSNADWV 239
 DB 330 EEKEVSTSSSYIEKYESSEDLSSSEQOQVPPVSKSLKQMSA--PAKPVNPIQTA--- 384
 OY 240 WLVRHLKLCMELCNNTYQMHLDLNCMEEPPIFGDPFLLPSFOSESSTPTGCG--FS 297
 DB 385 -----QQQ-----PAMVNISGPKKSGEPLS 406
 OY 298 GKTPSEDDRSQSR--EHMGESLSLKAGGDLPL--PSPKVEK----- 338
 DB 407 PQTGYSRLRSQSEQVLAEBEGVVAADMDIIMQPPRRRLQEVASAGBELVEKSIAY 466
 OY 339 -----DPSRKKEMWENAGN-----KITTMADKTISS-- 364
 DB 467 QPSGVPQPSVSKTIKEISLGAESAPLESPPRONHFORLLKPYQQVSAFERVISMDP 526
 OY 365 KLMEYKRRKO--OHNLSAPFEVKVKEKPEPLGPRQD-----SPLDORPH----- 410
 DB 527 KLTPQFORMRKQSGQSSSTPESTAF-----GSTVDIMTLTLLSPMLQAPQEVPLE 580
 OY 411 -----LMDQGMHRSFA-----GPELL----- 428
 DB 581 SETVAKMIKSTGOLSHKYSAPRLDSHQSPVPESTSAEGEELLSQPSVTPKFOPLMTQG 640
 OY 429 -----RQDKRRSGSTGSSLSVSVDAEAQIQAWNTVLTVLNQLQLP 473

Db 641 SIPSAMAPIYSPAPRMDSKP---LMSGALALESKSPRYSLQPMQS---TPFGQVSTPD 694
 QY 474 QTFAL-----QPAVFCISOLTC-HVTDIRVQA 502
 Db 695 HDPAAASWSPPIDPPTSRIPSOPLMRILAVKOPTCTELASVSQS 741

RESULT 8

ID 017596 PRELIMINARY; PRT; 2150 AA.
 AC 017596;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE C25D7.3 protein.
 GN C25D7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ainscough R.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT 'Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.';
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81039; CAB03774.1; -;
 DR InterPro: IPR000822; Znf_C2H2.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 KW DNA-binding; Metal-binding; Zinc-finger;
 SQ SEQUENCE 2150 AA; 249768 MW; E6A29F205124B589 CRC64;

Query Match 4.6%: Score 124; DB 5; Length 2150;
 Best Local Similarity 20.6%: Pred. No. 0.48; Indels 210; Gaps 24;
 Matches 114; Conservative 63; Mismatches 166;

QY 33 IIELPDEKNGHTKKSFSREIYVSL-----SHOVLQ-----NIYDILLEFVK 79
 Db 1177 IERLSEQLLGDOGSEDIISFEIYOVDLLESGEVQVQNSYTSRSSTSPESLVEDPEE 1236
 QY 80 GP-----SPGEKTIQ-VPEAKLAGLRYISMOMLAVIFDLLDYRTAREEDTSPGLKC 133
 Db 1237 HPEQLPVASEKANNQIVPEVEGVVPTNOQ-----EENVYSEG--P 1279
 QY 134 LKKVSGIGGAANLYRGSAMSFNIYFHALYCAVLTNO-----ETITAQ----- 177
 Db 1280 TLOGGSIPSSSHY-----TYDELGTSPGPEATEETVAEESPKKSGKT 1336
 QY 178 -----YKRVLFEDDERSTDSSQCSSEDEIFETAYOVSPPRGKEKROWRAMPL-L 228
 Db 1327 TRGPRKKVKNLKKRIQPRGQKEEAHPE--VVEEQOYVEPEVHEPVPAPAAQL 1385
 QY 229 SVQPVSNADNWLKRLHKLKLMELCNNTIOMHDLNEMCEP-----PIKRGDFLLPS 283
 Db 1386 ETEPI-----FOOIEPDKVFPEITEA-----LPL 1410
 QY 284 FOSSESTPTGFGSGKETPSDD-----RQSGREHMG 315
 Db 1411 FETSPVAPRCNIPSRASHSDDDVOYISSETDPNGPINLYEVOVONDKLTAYOYSTEEILG 1470
 QY 316 ESLSLKAGG-----DLLPPSPKVEKDPKSKKEMMEN--AGNKITYMAA 359
 Db 1471 EYGLDAGAPSPSEIYVHDEVLODPAPKSKKRGRRKKTTPHIKAKARVFT--- 1527
 QY 360 DKITSKLMTE-----YKKKQOHNLSA-----FPEKVEVEKKEGP-----LGR 398
 Db 1528 --SISK--TEEIELAPPTOOSRKRMANVSSSEATATRAKORAKAVEEPNDVSRLVLP 1583

QY 399 GODSPILORPOHLMDOGMRHSPSAGPELLROKRRPSSGSSLSYS----- 446
 Db 1584 PEDHETERPGHGEBSFETPSLTGR-----STASSVKTSSKRLFLSKNNP 1632
 QY 447 VRDAEQIOAMTN 459
 Db 1633 VPRMRIOQAGTN 1645

RESULT 9

ID 0961X4 PRELIMINARY; PRT; 809 AA.
 AC 0961X4;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE GH01088P.
 GN CG10732.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY047499; AAK77231.1; -;
 DR FLYbase; FBgn0036365; CG10732.
 SQ SEQUENCE 809 AA; 89641 MW; 9B62B41AFA15FA81 CRC64;

Query Match 4.5%: Score 122; DB 5; Length 809;
 Best Local Similarity 21.1%: Pred. No. 0.18; Indels 140; Gaps 23;
 Matches 105; Conservative 72; Mismatches 181;

QY 13 DTQCSPTPNPNFHAQSCQILIELP-PDEKNGHTKKSFSREIYVSLSHOVLQNLQYL 71
 Db 23 DVALSMDNNAQSIQSHSVLETPPTPTP-----SFDELOQRLEASNRMOHLHD 74
 QY 72 -----ILLEFVK-----GPSGEKTIQVPE--AKLAGFLR- 101
 Db 75 EQAKLQIONLAKTHLNEMERLRQADSLPHNVNGGEAPKYESVQOVDDVASLVGRKN 134
 QY 102 ---YISMON-LAVIF-----DLL-----LDYRTAREEDTSPGLCKLKKVSGIG 143
 Db 135 LTAFTIHONELSTVLGDDGPETLADQELQKLSLRTQRE-----DMRNLVDELNSINR 189
 QY 144 AANLYROSAMSFNIYFHALYCAVLTNOETTAQVKVLFEDDERSTDSSQCSSEDEDI 203
 Db 190 AA---RETA-----RVIKKEETPIPPPKP----- 211
 QY 204 FEETAYOVSPPRGKEKROWRAMPLLSVOVSNADNWLKRLHKLKLMELCNNTIOMHDL 263
 Db 212 -EEAAPAPKERVVPEVORVPIIR-OEANA-----AORLHQAOM-----INOKTADI 260
 QY 264 ENCMEEPPPIFKGDDFFILPFSFOSSESTPTGFGSGKETPSDDDSQSRHEHGESLSKAG 323
 Db 261 EALKAQMARLKG---MLYTSQIDESTPSMG--STLERRSERISVERELPAETLAQRYEA 315
 QY 324 GGDILLPPSPKVEKDKSKKEMENAGNKIYMAADKTJSLKLTETKRRKQOHNLSAP 383
 Db 316 LNDV-----TSELRAEASIAQK-----RRIILA-LKAEIERKQQAQAAVOM 357
 QY 384 KEVVEKKEGPELGRGDSPLORPOHLMDOGMRHSSAGP--ELRROKRRPSSGSGS 441
 Db 358 GDDALKRSILTPTP---TPMRORHTEEDSEVDNSLQATPTPEQLADELSOCERLRT 413

OY 442 SLASVDAEAOIOWTN 459
DB 414 EYKORLEFORVYASNN 431

RESULT 10
OY 09VU76 PRELIMINARY; PRT: 1657 AA.
AC 09VU76;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CG10732 protein.
CN CG10732.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Ffankoch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisl M., Kalush F., Karen G.H., Ke Z., Kesterson J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003538; AAF49814.2;
DR FlyBase: FBgn0036365; CG10732.
SQ SEQUENCE 1657 AA; 183992 MW; FF55F036A233ECBD CRC64;

Query Match 4.5%; Score 122; DB 5; Length 1657;
Best Local Similarity 21.1%; Pred. No. 0.49;
Matches 105; Conservative 72; Mismatches 181; Indels 140; Gaps 23;

OY 13 DTQCSPTPNNDHAQSCQILIEIP-PDEKPNHGHTKSVSFRELTVLSLHQLVLDLND 71
DB 301 DVALSEMSDNRASIOSLHSLVLETPRDPDTP-----SFDELQOHLSEASNNRMQHLHD 352

OY 72 -----ILLEEFVK-----GPSGGEKTIQVPE--AKIAGFLR- 101
DB 353 EQAKLQIQNLAKTHLNEMERLROQADSLPHNVNGEAPKYESVQVDDMASLVGRKKN 412

OY 102 ---YISMON-LAVIF-----DL------LDSYTAREFDTSPIGLKCLKKVSIGG 143
DB 413 LTAFLHONNELSTVLGDDGPEILAEQEOALOEKLESLERFORE-----DMRNLYDELNSINR 467

OY 144 AANLYROGAMSEFNIFYHALVCAVLINQETITAEQVKVLFEDDERSTSSQCSSEDDI 203
DB 468 AA-----RETA-----RIYKKEETPIPPPP----- 489

OY 204 FEETAQVSPPRGKGRQARAPPELLSVQVSNADWMLVKRLKLMELCNNTYOMHLDL 263
DB 490 -EEAAPAPKBRVAVVEYQVNVPIIR--QEAANA-----AQRALHQAAM-----INQKTADI 538

OY 264 ENCMEEPIFFGDDFFILPSFQSSSPSTPGFSGKTERPSDDDSQSEHNGESLSKAG 323
DB 539 EALKRAQMARLKG---MLTVTSQIESSTPSWG--STLERKSEERTSVRELPAETIAQRYFA 593

OY 324 GGDLLPSPPKVEKDKSPKKEWENAGNKIYTAADKTIKLTETKKRQOHNLSAFP 383
DB 594 LNDV-----TSLRRAEASLQKE-----RDRILA-LKAEIERRQQAANAAYOM 635

OY 384 KEVVEKKGEPLGPRGODSPILORPQHLMDOGMHSEFAGP--ELLRODKRPPSGSTGS 441
DB 636 GDDALKRLSLPTPT-----TPMRQMTREEDSEVDNSLQATPTKEQLRDELRSOCERLRT 691

OY 442 SLASVDAEAOIOWTN 459
DB 692 EYKORLEFORVYASNN 709

RESULT 11
OY 08FE17 PRELIMINARY; PRT: 1433 AA.
AC 08FE17;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4931;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB533; TRANSPOSON=LTR-RETROTRANSPOSON TS04;
RA Neugebelle C., Feldmann H., Bon E., Galliardin C., Casaregola S.;
RT "Genomic evolution of the LTR-retrotransposons in hemiascomycetous
yeasts."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ439550; CAD29535.1;
FT NON TER 1
SQ SEQUENCE 1433 AA; 164312 MW; 87C4075851F039BA CRC64;

Query Match 4.3%; Score 116.5; DB 3; Length 1433;
Best Local Similarity 19.9%; Pred. No. 1.2;
Matches 78; Conservative 65; Mismatches 142; Indels 107; Gaps 16;

OY 167 LTNQETITABQVKVLFEDDERSTSSQCSSEDDI FEETAQVSPPRGKGRQARAP 226
DB 584 VTNDVDVSELOSKENITTESE--TNEINKPSNTDEVYEENVRIPPAIQDN----- 633

OY 227 LLSVQP---VSNAQMWLVKRLKLMELCNNTYOMHLDLNECMEEPIFFGDDFFILPS 283
DB 634 LVGSKXTIININENI--ASRMQKNISGENETNYKELDDSDSCS----- 675

OY 284 FQSESTPSTPGFSGKTERPSDDDSQSEHNGESLSKAGGDDLLPSPKVEKKDPSRK 343
DB 676 -LHSDTNDSVYITTSKDNLTDDKDLQSQOELFEKVS-----DPEVLBEHMKIERDEVSON 729

QY 344 KEMMENAG-----NKITYMAADKTIISKLMTEYKKRKOQHNL-----SAPREVK 387
DB 730 SDNTSGVGFQFQPSIYTSSTNDQTN--DDYSTDKSHHLPYVNMNDNDQYDKPKN 787
QY 388 VEKGE---PLGPRGQDSPLOPQHLMDO-----414
DB 788 EKSNNNSDISISPGNNEDELQO---LVDSNKAKEKODATLESSAITDEPIEENDPANKA 843
QY 415 GOMHNSFAGPELLROKRRPSSSTGSSLSVSDAEAOQAMTNMVLTVNOQIILPDQ 474
DB 844 GFLKAFNS---LKKKKRPRLENKT--SPNDTAKRDNKROR-----NITKILPDN 889
QY 475 TPTALQPAVFPC--ISQLTCHVTDIRVQAVRE 505
DB 890 TETSSAPRIKTIYYNEAISRNADLKEKHAYKE 921

RESULT 12

Q8WZV3 PRELIMINARY: PRT: 2309 AA.
AC O8WZV3:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Conserved hypothetical protein.
GN B7N14.260.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohnsels J., Brandt P., Fairmann B., Holland R.,
RA Wakakura G., Mewes H.W., Mannhaupt G.,
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL669986; CAD21060.1;
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_8.
DR PRINTS: PRO0019; LEURICHRPT.
DR SMART: SM00370; LRR_3.
DR SMART: SM00369; LRR_Typ_8.
KW Hypothetical protein.
SQ SEQUENCE 2309 AA; 252994 MW; 583ACFE6D3A12B67C CRC64;

Query Match 4.3%: Score 115; DB 3; Length 2309;
Best Local Similarity 18.8%; Pred. No. 3;
Matches 93; Conservative 71; Mismatches 174; Indels 164; Gaps 23;

QY 66 LQNLVDLLEEFVNGPSPGEKTIQVPEAKLGF--LRTISMQMLAVFDLLDSYRTAR 123
DB 491 LNDISDLNPEESLQASTPK-----RPPSVKLEMDVDENHVAQVSGHSDSI--SPRTS 543
QY 124 EEDTSPGKCLLKRV---GIGGANLYROSAMSFNIYFHALVCALVLTNOETITAEQV 179
DB 544 AVPRIPULSVSTKRVASQAGSPLKIFP-----YDTFTSETIQ 584
QY 180 KVL--FEEDDE--RTDSSQOCSSDE--DIFETAVQSP-----RGKEKQRQRARMP 227
DB 585 RRLSQYHDDTPGEQASSGQGTGDESPVHVSYMDSGYISPKPTYTRKQETGQKRKASVAV 644
QY 228 -----LSVQPVSNADWWMLVLRKLKLMELCNNTYQMHLDLEN-----265
DB 645 NQFGAVLDGEYFQDDPSYQSVGAS-----IEGKKNQGRPD 683
QY 266 CMEBP--PIKGDFFILPSTQSSSTPSTGFGS-----KEPSEDDRQSRKHEMGES 317
DB 684 SQEPRIPIVFD---FSRPSRSRSEADAGDDADASNSVVRKRORPADSASSKRHSRMSS 739

QY 318 LSLKAGGDLPLP-----PSPKVE-----KQPSRKE-----WMEN 349
DB 740 MAKKRAAFTHIPATFTGPPSKRRSDPADEFKRRPTSPSKDTPKRRRTLHESIIEGLD 799
QY 350 AGNKITYMAADKTIISKLMTEYKKRKOQHNLSAPREYKVEK---KGEPLGPRGQDSPLO 406
DB 800 ISPSPEQSMKLSYQOQIALEKRR-----KEYNAEKHRQVEGTPSGKRDVPRPR 850
QY 407 RPQHLMQOGOMRHSFSGPELLRODKRPRS-----GSTGSSLY--445
DB 851 TPTPLQSSLOKDKCKGRPTELITRPPSRTPRSSKASGKGRTPKENGSGPSAMTDR 910
QY 446 --SVRDAEAOQAMTNMVLTVLQ 467
DB 911 KPSTKTEDELEA--NKMAMIRK 932

RESULT 13

Q8TE71 PRELIMINARY: PRT: 1077 AA.
AC Q8TE71:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE EGG1L.
GN EGG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA AebiJna1 W., Miller J.L.; EMBL/GenBank/DBJ databases.
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY074490; AAL71549.1;
SQ SEQUENCE 1077 AA; 120974 MW; 2B88BF3C47D032D6 CRC64;

Query Match 4.2%: Score 114; DB 4; Length 1077;
Best Local Similarity 18.0%; Pred. No. 1.2;
Matches 106; Conservative 91; Mismatches 223; Indels 170; Gaps 21;

QY 22 NNPDAQSCQ-----LLELPPEDEK--NCHTKKSVFREIVVLSHQVLLQNLVDIL 73
DB 172 HNLEFAKELQKTFSGLSLDLAKQKQRRHEMLKLEKKKKRTILQOVYVLQNLQEN 231
QY 74 LEEFVNGPSPGEKTIQVPEAKLGFRT-----SMQNLAVFDLLDS 118
DB 232 VQKDFKG---GLNCAVYLPSEKELDYLIKFSKLTCPERNESLSVEDOMQSSLYFWDLEG 288
QY 119 YRTAREPDTSPGLCKLKRVSGIGGANLYROSAMSFNIYFHALVCALVLTNOETITAEQV 178
DB 289 SEKAVVGTTKYHLDLSKLNLC-----YESTIPVKNKAKEVPLEE-332
QY 179 KKVLFEDERSTDSQOCSSDEDEIFE--ETAQVSP-----212
DB 333 -EMLIQSKTKQOLSKTESVKSSESLMERAQOEIQRQFELNRYMTEVDYSKKQEBQPM 391
QY 213 -----PRGKEKQ-----WRA-----RMLLSVQPVSNADWWMLV 242
DB 392 ADVARRPMLPRMWMLEPPDQOEKQESFKSMEASGKHQVESKPAVSL-----440
QY 243 KRLKLMELCNNTYQMHLDLENCEEPPIKQDPFLPSTQSSSTPSTGFG-----S 297
DB 441 -----QRKQDTSKRLSTLPEDQKQEIISKRP--SPSQMKQDTPKSAAGVVOEQK 489
QY 298 GKETP-----SEDDRSQSRHEMGESL--SLKAGGDLPLPSPKVEKKRPSRKEMW 347
DB 490 KQETPKLMPVOLQEKQDQKQKTPKSWTSMQSSQNTTSMWTPHCEEDDSQOPEPKSME 549
QY 348 ENAGNKITYMAADKTIISKLMTEYKKRKOQHNLSAPREYKVEKGEPL--GPRGQDSP 404
DB 550 NNVESOKHSLTSQSSISPSKSWGAVATASLIPNDQLPRKLTNEPRKDVPRVHQPVQSSSTL 609

OY 405 LARP-----OHLMDOGOMHSHFSAGPELLROD-----RPRSGTSSLSVSVRD 449
 DB 610 PKDVLRRKELQDLOTCGTCNFMQ--ESVLDPDKPSPALPTSQPSPATPSPSVASKRON 668
 OY 450 AEAQ-----IQAMTNVLTVLNQIOLPQOT-----FTALQPAVF 484
 DB 669 LSSSDLEPQLQATSSPVTCSSNACLVTYDQASSGSETEPMTSETPEVF 718

RESULT 14

092529 PRELIMINARY: PRT: 594 AA.
 AC 092529;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE P64 isoform of N-SHC (P52 isoform of N-SHC).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN:
 RX MEDLINE-96404502; PubMed-8808684;
 RA Nakamura T., Sanokawa R., Sasaki Y., Ayusawa D., Oishi M., Mori N.;
 RT "N-Shc: a neural-specific adaptor molecule that mediates signaling
 from neurotrophin/Trk to Ras/MAPK pathway.";
 RL Oncogene 13:1111-1121(1996).
 DR EMBL: D84361; BAA12322.1; -;
 DR EMBL: D84361; BAA12323.1; -;
 DR HSSP: P29353; ISHC.
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00640; PID; 1.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PRO0629; SHCIPDOMAIN.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00462; PTB; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS01179; PID; 1.
 DR PROSITE: PS00001; SH2; 1.
 SQ SEQUENCE 594 AA: 64056 MW: 047AD9D08002E41BF CRC64;

Query Match 4.1%; Score 111; DB 4; Length 594;

Best Local Similarity 18.5%; Pred. No. 0.95;

Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

OY 83 GPEKTIQVPE-----AKLAGFLRYI-SMONLAVIFDILLDSYRTARE 124
 DB 129 PGDEPLRRPRGARNASDQVLGPGVTVYVVKLGCEIVLRSMKSLDFSTRQITREAIKRV 188
 OY 125 FDTSPGLKCLKK-----VSGIGGANL-YROSAMSFNYFHALVCAVLTNOETITAE 176
 DB 189 CEAVPGAGAKGKKRPRSKMLSTLQFAGMSISLTSTASLNLRTPDSQIITANH 248
 OY 177 QVKVLEFD-----DERSTDSQOCSSEDEDFEETAQVSPRRKEK 218
 DB 249 HMRISIFASGDDPTTDVAVVARDPVNRACHILECCDGLADVIGSIGAFELRFKQY 308
 OY 219 ROMRABMPLSVQPVSNADWVLVRLKLCMELCNNTIQMHLENCMEBRPFKGRPF 278
 DB 309 LQCFPTKIPALH-DRMQSLDERP-----TEEDGDSHPHYNSIPS 347
 OY 279 -----FILPSFQSESTPSTGSGFKETPSEDDRSQREHMGESLS-----LKAG 324
 DB 348 KMPRGGLDRLRLRPRAAPRTAQFAGKEGYUGR-----HLDGTFEDMQQTRPLRGS 402
 OY 325 GDLLPSPKVEKDPKSKKEMENAGNKIYTMAADTKISLMTYEKKRKQOHLSAPK 384
 DB 403 SDIVSTEGKLVHA-PTGEAPTYYNT-QQIPQAMPAAVSS--AESPRKDLFMKPFED 458

OY 385 EVKVEKKG-----EPLGPRGDSPLLRPQ-HLMDGOMHSHFSAGPELLROD 431
 DB 459 ALKNQPLGPLYSKAASVEICISPVSPRAPDAKMLBELQAEITYQSEMRSKEAG--LLEKD 516
 OY 432 ----KRRSSSTGSSLSVSDAEQAOAMTNVLTVLNQIOLPQOTFTLQPAVFPPI 487
 DB 517 GDFLVKRSTNPGSFVLGMHNGQAK-----HLLLVPEGTIRT-KQVVPDSI 563

RESULT 15

08TAP2 PRELIMINARY: PRT: 594 AA.
 ID 08TAP2;
 AC 08TAP2;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Neuronal Shc.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN:
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC026314; AAH26314.1; -;
 DR EMBL: BC026314; AAH26314.1; -;
 SQ SEQUENCE 594 AA: 64042 MW: FE484556345FF6C5 CRC64;

Query Match 4.1%; Score 111; DB 4; Length 594;

Best Local Similarity 18.5%; Pred. No. 0.95;

Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

OY 83 GPEKTIQVPE-----AKLAGFLRYI-SMONLAVIFDILLDSYRTARE 124
 DB 129 PGDEPLRRPRGARNASDQVLGPGVTVYVVKLGCEIVLRSMKSLDFSTRQITREAIKRV 188
 OY 125 FDTSPGLKCLKK-----VSGIGGANL-YROSAMSFNYFHALVCAVLTNOETITAE 176
 DB 189 CEAVPGAGAKGKKRPRSKMLSTLQFAGMSISLTSTASLNLRTPDSQIITANH 248
 OY 177 QVKVLEFD-----DERSTDSQOCSSEDEDFEETAQVSPRRKEK 218
 DB 249 HMRISIFASGDDPTTDVAVVARDPVNRACHILECCDGLADVIGSIGAFELRFKQY 308
 OY 219 ROMRABMPLSVQPVSNADWVLVRLKLCMELCNNTIQMHLENCMEBRPFKGRPF 278
 DB 309 LQCFPTKIPALH-DRMQSLDERP-----TEEDGDSHPHYNSIPS 347
 OY 279 -----FILPSFQSESTPSTGSGFKETPSEDDRSQREHMGESLS-----LKAG 324
 DB 348 KMPRGGLDRLRLRPRAAPRTAQFAGKEGYUGR-----HLDGTFEDMQQTRPLRGS 402
 OY 325 GDLLPSPKVEKDPKSKKEMENAGNKIYTMAADTKISLMTYEKKRKQOHLSAPK 384
 DB 403 SDIVSTEGKLVHA-PTGEAPTYYNT-QQIPQAMPAAVSS--AESPRKDLFMKPFED 458
 OY 385 EVKVEKKG-----EPLGPRGDSPLLRPQ-HLMDGOMHSHFSAGPELLROD 431
 DB 459 ALKNQPLGPLYSKAASVEICISPVSPRAPDAKMLBELQAEITYQSEMRSKEAG--LLEKD 516
 OY 432 ----KRRSSSTGSSLSVSDAEQAOAMTNVLTVLNQIOLPQOTFTLQPAVFPPI 487
 DB 517 GDFLVKRSTNPGSFVLGMHNGQAK-----HLLLVPEGTIRT-KQVVPDSI 563
 OY 488 SOLTCH 493
 DB 564 SHLINH 569

Wed Jun 18 09:54:38 2003

us-09-991-681-27.rspt

Page 9

Search completed: June 17, 2003, 11:59:43
Job time : 132.931 secs

THIS PAGE BLANK (USPTO)